DOCUMENT RESUME

ED 309 180 TM 013 628

AUTHOR	Anrig, Gregory R.; And Others
TITLE	What Is the Appropriate Role of Testing in the Teaching Profession? Proceedings of a Cooperative Conference (December 12, 1986).
INSTITUTION	National Education Association, Washington, D.C.
REPORT NO	ISBN-0-8106-1460-X
PUB DATE	87
NOTE	86p.
AVAILABLE FROM	National Education Association, Professional Library, P.O. Box 509, West Haven, CT 06516.
PUB TYPE	Collected Works - Conference Proceedings (021) -- Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Accountability; Educational Policy; *Educational Testing; Elementary Secondary Education; Teacher Education; *Teacher Evaluation; Test Construction; *Test Use
IDENTIFIERS	Educational Testing Service; National Education Association; *Teacher Competency Testing

ABSTRACT

Five papers given at a conference on the role of testing in the teaching profession are presented. The conference partners included the American Association of Colleges for Teacher Education, Educational Testing Service, Center for Fair and Open Testing, and National Education Association. The following issues are discussed: appropriate uses of testing within the teaching profession; who should control of the content, construction, and use of teacher tests; translation of public demands for teacher accountability into appropriate and equitable testing policies and procedures; and steps for preventing the misuse of teacher tests. The papers include: (1) "The NEA Perspective on the Role of Testing in the Profession" (Sharon P. Robinson); (2) "Teacher Testing in American Education: Useful But No Shortcut to Excellence" (Gregory R. Anrig); (3) "Testing Teachers: Strategies for Damage Control" (John G. Weiss); (4) "The Appropriate Role of Testing in the Teaching Profession" (Norene F. Daly); and (5) "Highlighting the Critical Issues" (Louis J. Rubin). Conference participants' observations and recommendations and panelists' reactions to the observations and recommendations are provided. Appropriate methods, times, and subjects for testing; the relation of testing to professional preparation and training; testing of minorities; and development and implementation of an agenda to implement testing policy are also discussed. The names and affiliations of the conference participants are listed. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED309180

What Is the Appropriate Role of Testing in the Teaching Profession?

Gregory R. Anrig
Norene F. Daly
Mary Hatwood Futrell
Sharon P. Robinson
Louis J. Rubin
John G. Weiss

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

G. FELTON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM013628

A National Education Association Publication

2

What Is the Appropriate Role of Testing in the Teaching Profession?

Proceedings of a Cooperative Conference

American Association of Colleges for Teacher Education
Center for Fair and Open Testing
Educational Testing Service
National Education Association

National Education Association
Washington, D.C.

Copyright © 1987
National Education Association of the United States

Library of Congress Cataloging-in-Publication Data

What is the appropriate role of testing in the teaching profession?

1. Examinations—United States—Congresses.
2. Teachers—Certification—United States—Congresses.
3. Teachers—United States—Rating of—Congresses.
I. American Association of Colleges for Teacher Education. II. National Education Association of the United States.
LB1762.W48 1987 379.1'57 87-22056
ISBN 0-8106-1460-X

CONTENTS

Preface, by Mary Hatwood Futrell 5

I. Defining the Appropriate Role of Testing in the Teaching Profession

The NEA Perspective on the Role of Testing in the Profession ... 9
Sharon P. Robinson (NEA)

Teacher Testing in American Education:
Useful But No Shortcut to Excellence 21
Gregory R. Anrig (ETS)

Testing Teachers: Strategies for Damage Control 34
John G. Weiss (FairTest)

The Appropriate Role of Testing in the Teaching Profession 49
Norene F. Daly (AACTE)

II. Reconsidering the Appropriate Role of Testing in the Teaching Profession

Highlighting the Critical Issues 63
Louis J. Rubin (Conference Facilitator)

Participants' Observations and Recommendations 67

Panelists' Reactions to Participants' Observations and Recommendations 72

Epilogue 77
Louis J. Rubin

Conference Participants 78

PREFACE

On December 12, 1986, the National Education Association (NEA) hosted a one-day invitational conference to consider one of the most complex and controversial questions facing the profession today: What is the appropriate role of testing in the teaching profession? The reason for the conference lay in the belief that there is considerable confusion, lack of clarity, indecision, and anxiety surrounding this issue. The NEA decided to address it by gathering together, in one room, representatives of the four principal actors in the teacher testing drama—teacher educators, test developers, consumer advocates, and teachers themselves.

The conference partners—the American Association of Colleges for Teacher Education (AACTE), the Educational Testing Service (ETS), the Center for Fair and Open Testing (FairTest), and the NEA—have different views of testing in the profession. But they did agree to sit down together to examine this critical question in which they all have a stake.

The hopes for the conference were simple: that the day would provide some new insights for all participants, that we could begin work on a consensus concerning testing's role in the profession, and that we could pave the way for continued exchange.

The day's structure was straightforward. Each organizational partner presented its own perspective on the matter of teacher competency testing in response to four questions posed by the conference planners:

1. What are the appropriate uses of testing within the teaching profession?
2. Who should control the content, construction, and use of teacher tests?
3. How can public demands for teacher accountability be translated into appropriate and equitable testing policies and procedures?
4. What steps can be taken to prevent the misuse of teacher tests?

All the participants considered the perspectives offered by each organization in small group discussions and developed recommendations for tackling a range of issues related to testing's appropriate role in the teaching profession.

These proceedings contain the four major presentations, each with a follow-up exchange of questions and answers; a synthesis of the presentations, which motivated the group discussions; a list of observations and recommendations culled from the group deliberations; reactions to the group reports by the panelists; and a list of the conference participants.

The NEA expresses its sincere thanks to all who joined in this endeavor. Now we must build upon the initiative.

Mary Hatwood Futrell
President, NEA

I. Defining the Appropriate Role of Testing in the Teaching Profession

THE NEA PERSPECTIVE ON THE ROLE OF TESTING IN THE PROFESSION

Sharon P. Robinson, Director
Instruction and Professional Development
National Education Association

What is the appropriate role of testing within the teaching profession? Let me say unequivocally that it must be whatever *the profession* determines it to be. And by profession I mean practicing classroom teachers who teach in settings ranging from preschool to graduate school. As other professions have done, teachers must decide what testing will be in their own profession. As other professions have done, teachers must answer the why, how, when, and what of professional testing. In short, teachers must govern both the content of professional examinations of teaching as well as the procedures used to administer those assessments.

If that seems a hardline position, let me add that teachers cannot determine the appropriate role of testing within their profession without the assistance of other groups. We must address difficult questions related to the content of professional examinations in teaching, the effect of such exams on minority recruitment in the profession, and the structure of the exams themselves. Teachers will be able to resolve these issues by working closely with deans of teacher education and other higher education faculty members, psychometricians and test developers, and watchdogs of the testing process.

Let me share with you the NEA's responses to a series of challenging questions related to testing within the profession that were posed to this panel by the conference planning group:

1. What are the appropriate uses of testing within the teaching profession?
2. Who should control the content, construction, and use of teacher tests?
3. How can public demands for teacher accountability be translated into appropriate and equitable testing policies and procedures?
4. What steps can be taken to prevent the misuse of teacher tests?

I will add one more question to the list:

5. How do we know what to test?

In the course of answering these questions, I hope to convey the NEA's views on the need to professionalize teaching, on professionalization and testing, and on professional self-governance. First, the NEA's position on teacher testing must be made clear, for that position provides a foundation for the related views.

The NEA believes that every classroom must be staffed with competent, licensed, and properly assigned teachers. There can be no room for compromise through emergency certification, teacher misassignment, or less than rigorous preparation. In this context, there is a definite, legitimate, and even valuable role for the fair and judicious use of testing. Just as no law graduate is allowed to practice law without successfully completing the bar exam, no teaching graduate should be entrusted with the instruction of America's public school students without demonstrating mastery of subject matter and professional skills (Futrell & Robinson, 1986). We believe that tests are an important way to measure that mastery, particularly when they are part of a broader assessment context. I will elaborate on that context later, but now will turn to what is perhaps the most important and basic question underlying our discussion: How do we know what to test?

WHAT TO TEST

Knowing what to test relates to a longstanding debate among teachers, teacher trainers, researchers, and measurement specialists about what constitutes the knowledge base of teaching. Is there a codified knowledge base of teaching on which we can all agree? Yes and no. If pushed to do so, I think we could reach agreement on the skills and knowledge that comprise at least the beginnings of the knowledge base underpinning teaching, but neither teachers, teacher educators, nor test developers have really been pushed to that point. Until now. The prod for action is Carnegie's national board initiative.

Shulman and Sykes (1986) have written that the national board "rests on the claim that distinctive knowledge for teaching exists, that teachers use their knowledge in a distinctive way, that current approaches to the assessment of teaching competence do not tap adequately or credibly into that base, and that a new approach to assessment can be developed to remedy those problems" (p. 5). They go on to offer a definition of the knowledge base for teaching. It is, they suggest, "that body of

understanding and skill, of dispositions and values, of character and performance, that together underlie the capacity to teach" (p. 5).

The categories of knowledge that teachers need to master are "general/liberal education, including basic skills of reading, math, writing, and reasoning; content knowledge in the domains in which teaching will occur; content-specific pedagogical knowledge; general knowledge of pedagogical principles and practice; curricular knowledge; understanding of student diversity and individual differences; performance skills (including voice, manner, poise); and foundations of professional understanding (including history and policy, philosophy and psychology, cultural and cross-cultural factors, and professional ethics)" (Shulman & Sykes, 1986, p. 7).

This list rests on a conception of teaching as the transmission of knowledge and understanding through instructions. Consequently, it is limited. But this limitation is both acceptable and necessary. Let me say that another way: The NEA finds this delimitation of teaching reasonable.

First, we do not believe any examination system will ever assess all that teachers must know and be able to do in order to teach. Although it has been noted that the task environment of the classroom is more complex than that faced by a physician during a diagnostic examination (Shulman, 1984), teachers get precious little credit for the job they do. In the final analysis, they are in the unenviable position of doing a cognitively and conceptually complex job—and making it look easy.

Second, on purely practical grounds, let me suggest than an examination that truly assesses mastery of the categories of knowledge laid out by Shulman and Sykes (1986) will be a noticeable improvement over current practice. We all know the limitations of current tests of teaching skill and knowledge. As Linda Darling-Hammond (1986) has observed, "Current teacher tests are limited, as other professional tests are not, by the exigencies of a multiple-choice format, by the lack of definition of the basis for the knowledge base and how it can be well-demonstrated, and by lack of recognition that complex reasoning and judgment abilities are fundamental to both subject matter knowledge and pedagogical knowledge" (p. 9).

Probably we all would agree with that assessment, as probably we all would acknowledge that teaching is definitely more than transmitting knowledge and understanding through instruction. But it is also at least that, and we must bear this fact in mind as we seek to codify a knowledge base for teaching and develop ways to test mastery of that knowledge base.

Finally, it needs to be said emphatically that teachers must steer the codification of teaching's knowledge base. First, it is their professional responsibility to do so. Second, they are well-prepared to do so. Third, it is more than ever in their self-interest to do so.

TESTING, ACCOUNTABILITY, AND TEST MISUSE

These points lead to two related questions: How can public demands for accountability be translated into appropriate and equitable testing policies and procedures? What steps can be taken to prevent the misuse of teacher tests? The public's demand for accountability *must* be translated into appropriate and equitable testing practices within teaching to prevent the kind of misuse of testing we have seen to date and possible future abuses. Nothing less than the future of teaching is at stake. We have already experienced the devastating impact of ill-conceived, politically driven state-level testing mandates on the breadth and depth of the teaching force. These tests are particularly worrisome at the precertification level, where their number is increasing rapidly.

We all would agree that it is foolish and unfair to expect a high rate of success on precertification tests from students and institutions of higher education with inadequate resources. Yet that is exactly what we are asking of disadvantaged minority students and historically minority institutions. And they are not meeting these new demands. As Lloyd Hackley (1985) has stated, "The drive toward improved educational standards is catching unprepared thousands of persons who, through no fault of their own, are ill-equipped for higher standards and greater rigor."

We know that, if they prevail, current testing practices will contribute to the demise of the minority teacher. Furthermore, this will happen just when the number of minority students in the schools is increasing so dramatically that Hispanic and Black students, and students from other minority groups, will soon constitute the majority of our school-age population (Hodgkinson, 1986).

All the groups represented in this conference—teachers, teacher educators, test developers, and those who monitor the testing industry—must challenge the politics of test misuse with its focus on sorting and selecting. Current state-mandated testing for teachers merely puts a good face on what is politically expedient and politically acceptable. As a consequence, the tests merely validate what we already know based on resource allocation and access to learning opportunities. We know the only way to improve minority students' precertification test scores is through vast improvements in the education these students receive in

college and earlier, but politically driven and instituted testing mandates are far easier and cheaper. They meet public demands for accountability, but they fail to get at the root of an illness of which poor test scores are only the most visible symptom.

The NEA believes that it is in teachers' self-interest and that it is our professional responsibility to change the public's and politicians' naive and dangerous notions of accountability through testing's misuse in the form of legislated testing. In short, accountability itself must be defined by teachers and not by governors and state legislators. How teachers accomplish this redefinition hinges on control of the testing process. This, in turn, relates to control of licensure and certification, and ultimately, to professional self-governance.

This brings us to two other questions: Who should control the content, construction, and use of teacher tests? What are the appropriate uses of testing within the teaching profession?

THE CONTROL AND USE OF TESTING IN TEACHING

It is the NEA's view that the profession—classroom teachers—must play a dominant role in matters related to the content and use of teacher tests as a way to promote professionally developed and enforced standards. The actual construction of the tests is best left to experts at ETS and elsewhere. It has been pointed out, for example, that teacher tests are now developed by testing companies who carry on their test construction without a clear client. That is, there is no professional client for teacher testing as there is in other professions. The National Board of Medical Examiners and the American Bar Association, for example, serve as professional clients for testing companies who are in their employ, and the final decisions about content and substantive standards are made by those professional bodies (Darling-Hammond, 1986). I submit to you that the teaching profession needs a comparable body. The national board should begin to serve this function. Teachers, through delegated licensing authority at the state level, can require the application of professional examinations governed by their national board.

We believe that leadership from teacher educators is needed to direct assessment processes used in professional preparation and that leadership from teachers is needed to direct assessment required for entry into professional practice. One way of exerting that leadership is simply by advancing a model of the appropriate uses of assessment for (*a*) degree completion, (*b*) admission to practice, and (*c*) evaluation of practice.

Assessment for Degree Completion

We believe that Assessment for Degree Completion is the province of institutions of higher education, so all the deans and teacher education faculty members in the audience should listen with a critical ear to our suggestions.

Assessment for Degree Completion should assure colleges and universities, and the teaching profession, that a candidate has successfully completed an institution's degree program. Assessment should include a broad range of curriculum-embedded procedures to diagnose and measure students' knowledge, skills, and performance. But tests would be only one element of a comprehensive assessment plan developed and overseen by the faculty of the institution. Furthermore, we believe that the assessment plan should emphasize institutionally developed and curriculum-related assessment rather than standardized testing.

We recommend that institutions of higher education assess all students, including prospective teachers, at two critical points—the beginning of the freshman year and the end of the sophomore year. That is, we recommend diagnostic testing of all college freshmen in basic skills and general academic areas, regardless of the majors they pursue. These test results would be used by the faculty to advise students about course selection, course sequencing, and remediation, if needed, during the first two years of college. The institution would retain discretion to determine the uses of such criteria as college entrance exam scores and high school grade point average in relation to this freshmen testing.

We then recommend basic skills and general education assessment at the end of the sophomore year for all students. This would serve as a posttest of the first two years of college instruction for students. The results of such posttesting would advise the entry to various majors and the design of the upper-division curriculum. For institutions offering teacher preparation in their upper divisions, this assessment would serve as an entrance to the professional program. As such, the assessment would provide a basis for making one of three decisions about teacher candidates, depending on their scores: (*a*) entrance into teacher education, (*b*) entrance coupled with remediation, and (*c*) delayed entry. In this way, the test results would be used to determine the need for additional courses to strengthen a student's skills.

We recommend that students who demonstrate severe skill deficiencies be assigned to remedial classes within the institution's appropriate division, not to courses within the teacher education program. We believe, however, that these students should have the option of taking the

basic skills/general education test repeatedly to gain entry to the teacher preparation program. The institution would have the option of setting a limit on the number of retesting opportunities and thereby establishing the point at which unsuccessful performance on the test would exclude students from the teacher preparation program.

We also recommend that institutions of higher education assess and verify teacher candidates' knowledge of an academic content area, their professional knowledge, and their ability to perform in the classroom prior to granting these students a degree. More specifically, the NEA advocates performance evaluation of preservice teachers within an extended, supervised practicum. The practicum should be a series of classroom-based experiences in diverse settings and should (*a*) be initiated early in a candidate's academic career, (*b*) allow increasing independence and become more demanding as the candidate acquires additional pedagogical skill, and (*c*) culminate in a concentrated period of clinical experience.

Finally, we recommend a professional assessment of candidates' academic content areas and professional knowledge and skills prior to their exiting teacher preparation. To complete this assessment, colleges and universities could use available testing and assessment mechanisms or develop a variety of procedures (not just tests). That is, while this verification could be demonstrated on an exit examination, institutions could choose to base it on a range of indicators, such as performance in practicum experiences, grade point average, scores on faculty-developed examinations administered in teacher education courses, and completion of projects and papers.

We believe that students should have multiple opportunities to be assessed and to audit or retake courses within the teacher preparation program or other appropriate college/university divisions to gain the skills required for graduation and, ultimately, for successful classroom practice.

Our assumption, by the way, is that assessing students' mastery would fall to faculty members working within the confines of an accredited teacher education program. Accreditation's role is vital to assuring that a teacher preparation program is recognized by the profession as meeting the profession's standards for content and operation.

Assessment for Admittance to Practice

Once students have graduated from teacher training, they would be eligible for Assessment for Admittance to Practice. This assessment pro-

cess should be the province of the profession. It should place with the profession the responsibility for assuring that an individual is qualified for employment as a teacher. We believe that the profession would exercise this responsibility through assessment or validation procedures to be developed by the profession, with the advice and input of teacher educators, researchers, test developers, and measurement experts, and administered by state professional standards boards. These procedures would include comprehensive, multiple techniques and assessments developed and administered by the boards, which in turn would be composed of a majority of practicing teachers.

The NEA, as you know, is on record as an unrelenting advocate of rigorous certification and licensure procedures developed, administered, and monitored by state-based standards boards. The Association remains interested in the proposal of the Carnegie Task Force on Teaching as a Profession for creation of a National Board for Professional Teaching Standards. NEA President Mary Futrell and other NEA members serve on Carnegie's planning group for the national board.

However, let me clarify our position. First, the NEA has adopted a policy position which affirms our support of a voluntary national certification process. Second, we believe that national certification and a national certification board should under no circumstances usurp the authority of state-level professional standards boards. We recommend, and will work for, an integrated model. State boards would help enhance the prestige and visibility of the national board's standards, and would likely use or adapt those standards, but state boards would retain responsibility for the testing and licensing of teachers as well as for the administration of certification procedures.

Assessment of Practice

Once individuals are granted a license by the state, the public should have confidence that they are competent to teach. And the public should expect them to demonstrate their continued competence. This is accomplished through Assessment of Practice.

Assessment of Practice is the province of local school districts. First, it involves the ongoing, comprehensive evaluation of all practicing teachers by trained evaluators to determine their effectiveness in classroom instruction and interaction with students. Second, it entails the continued professional development of teachers to develop and refine their skills. At this point, testing ceases and evaluation takes over. We do not support the use of tests for recertification or relicensing.

In summary, then, on this issue of the appropriate role of testing within the profession, the NEA believes that assessment for degree completion is the responsibility of institutions of higher education and their teacher education faculties. Local school districts, in consultation with teachers, are responsible for assessment of practice. The profession is paramount when it comes to admittance to practice, for if teachers are to become self-governing, they must establish the criteria and standards for entry into their own profession and themselves must decide who shall and shall not join them as colleagues.

The NEA believes that, if used appropriately, testing can enhance the credibility and status of teaching as has been the case in other professions. Appropriate use hinges on—

1. Codifying a knowledge base for teaching
2. Advancing a model that defines how testing will and will not be used in the profession
3. Seeking professional control of certification and licensure
4. Demanding the end of politically expedient teacher testing policies.

References

Darling-Hammond, L. 1986 (April). Raising professional standards and teacher quality. Paper presented at the annual meeting of the American Educational Research Association.

Futrell, M.H., and Robinson, S.P. 1986. Testing teachers: An overview of NEA's position, policy and involvement. *Journal of Negro Education* 55 (No. 3).

Hackley, L.V. 1985. The decline in the number of black teachers can be reversed. *Educational Measurement: Issues and Practice* 5 (Fall): 17.

Hodgkinson, H.L. 1986. *Future search: A look at the present.* Washington, D.C.: National Education Association.

Lareau, A. 1986. A comparison of professional examinations in seven fields: Implications for the teaching profession. *Elementary School Journal* 86 (No. 4): 553-69.

Shulman, L.S. 1984. It's harder to teach in class than to be a physician. *Stanford School of Education News* 3 (Autumn).

Shulman, L.S., and Sykes, G. 1986 (May). A national board for teaching? In search of a bold standard. Paper prepared for the Task Force on Teaching as a Profession, Carnegie Forum on Education and the Economy.

* * *

Q. You argue that we ought to give teachers control over the content and administration of teacher competency testing, but won't the general public be concerned that teachers' control of their own destiny will not result in true accountability?

A. Such public concern would be a rather natural reaction; and if we encourage it, it may become the prevailing view. We have to help the public understand why it is in their interest to trust teachers with this kind of responsibility. Teachers are very critical and very concerned about the status and recognition of their profession. This is threatened when there are those within the teaching work force who cannot write, who cannot read, who are not competent to teach. We could argue that the system we currently have has produced this state of affairs. Why not give teachers a chance to do it right? They care about the standards and will be more uncompromising about those standards than others have been.

Q. You say that we must put an end to "politically expedient teacher testing." Do you mean to imply that those who administer tests have some political objective above and beyond determining competence—that testing is used to discriminate?

A. Yes. The test score is used to control the supply and demand of teachers. That the requisite cut score slides up and down the scale, depending on the exigencies of supply, bears on the economic well-being of teachers. When education can always meet its teacher supply requirements, there is no real need and no interest in transforming teaching into a competitive career choice. Further, the consequences of the disproportionate performance of minority students on these tests may not be an obvious intent, but it is an outcome that the politicians are willing to tolerate, a cost that they are prepared to pay. We must make that cost much more than they dare be prepared to pay.

Q. You say that we should not use tests to sort and select among teachers, that this is test misuse or test abuse. The logical question then is, if we don't use the tests to eliminate the weak, is there a point in testing in the first place? Should we not use the tests to get rid of practicing teachers who are incompetent?

A. The NEA's conception of the use of tests in the overall continuum of becoming a teacher, if you will, really ends when a person is admitted to practice. At that point, we feel, competence in practice should be determined through a cogent and real system of evaluation of performance, professional development, and constant oversight in terms of professional practice. And by constant oversight I do not mean routine administrative supervision. I mean commitment to quality practice that is supported and reinforced by a community which knows what that is and knows how to help all those in the enterprise achieve it.

Q. You speak about the need of codifying a knowledge base that teachers ought to have. You further suggest that the present tests leave a great deal to be desired. Do you really think that we can set down in print precisely what teachers ought to know and be able to do, with absolute finality? Isn't there always the case that someone who sees things differently might fail a test but be very acceptable?

A. I cannot answer that question unless I look at a different conception of how testing in teaching might be structured, in terms of using technologies and getting beyond the multiple-choice format. I suggest that we can answer the question of whether there is a knowledge base in teaching simply by daring to do so. It would not be the final answer. It would be an answer that we intend to perfect over time. If we do not discipline ourselves to get beyond our chicken-egg argument about whether there is or is not a knowledge base, the politicians are going to continue deciding that they want competency testing of literacy skills, and that's not the test that I want to have represent the teaching profession.

The test I envision would cost more than going to ETS or anybody else and saying give me a hundred of your best questions. If we want to help support teaching as a profession, and to help support quality education for students, we need to make the politicians understand that there is a big difference between professionally appropriate testing and testing to provide political answers to accountability questions. Furthermore, we must ultimately make politicians understand that students bear the cost of ill-conceived teacher testing.

Q. Doesn't your analogy to lawyers and doctors and their control over their licensing break down in light of the fact that they are self-selected by their clients and can be sued by them for malpractice? Second, assuming you rebut the implications of that question, would

you also argue that teachers should have the same kind of say in the structuring of preparation that, for example, lawyers have in the structuring of legal education and the certification of law schools, that is, determining whether or not institutions can be accredited?

A. First, I do not believe the arena of practice is all that relevant. Attorneys who serve in the public sector have to achieve the same standards and access to the profession as those who work in the private sector, and the same is true with doctors. Teachers already have a great deal of influence in establishing standards for teacher preparation through our historical participation in the National Council for Accreditation of Teacher Education (NCATE). The challenge suggested at this conference is that our next battleground may well be the institution's assessment program.

Second, in the interest of public accountability, we have to be clear about the system of credentialing in teaching. That system is licensure, which at the state level gives one legal permission to teach. I believe teachers can bring a great deal of guidance, a great deal of vigor and commitment to standards in a licensure procedure that relates directly to the public and satisfies the public's need for standards and accountability. I think we need to organize ourselves as a profession to offer a system of certification, if you will—or assessment for certification—which would admit one to the profession. Then we need to work politically to link certification to the licensure process. These steps would strengthen the system and create high standards and professional commitment as other professions have achieved.

TEACHER TESTING IN AMERICAN EDUCATION: USEFUL BUT NO SHORTCUT TO EXCELLENCE

Gregory R. Anrig, President
Educational Testing Service

The appropriate role of testing in the teaching profession needs to be viewed in the broader context of the unprecedented period of educational reform in which we find ourselves. The 1970s saw a burgeoning of state requirements for minimum competency testing of students. In the 1980s, the focus of state mandates has shifted to such testing for prospective teachers. Before 1980, only ten states had teacher testing requirements. Today, forty-four states do (Sandefur, 1986). What do these trends mean?

Before coming to the ETS in 1981, I served for more than eight years as Massachusetts Commissioner of Education. I had then—and still have—a healthy skepticism about standardized tests. At the same time, I respect their role when used and interpreted within their proper limits. The pressure for the testing of students and teachers is not coming from educators who are familiar with the limitations of such instruments, however; it is coming from the public and its elected representatives.

HOW WE GOT TO WHERE WE ARE

This pressure is the result of widespread dissatisfaction among taxpayers over what they perceived as falling educational standards in the 1970s. Ironically, educational trends during this period reflected prevailing community attitudes amid the Vietnam-driven social unrest of the time. (Remember the demands for "relevancy" and educational alternatives?) The tide turned abruptly, however, as the country's mood became conservative and its concern over jobs, the economy, and foreign competition heightened. Stimulated in part by extensive media coverage of the lengthy decline in college admission test scores, voter unrest led public officials to see school reform and demands for higher educational standards as *the* domestic political issue of the 1980s, both in the states and nationally. Because public officials viewed schools as a large part of the

standards problem, they looked for an external yardstick that could be used legislatively to force higher standards into the educational system. In the name of accountability, state after state enacted testing requirements, first for students and then for teachers.

Understanding this evolution leads to a basic conclusion: Testing is here and is not going to go away. Public education cannot thrive without taxpayer support, and taxpayer support—at least for the present—requires external assurance that students and teachers can meet at least some minimum standard of academic competency expected by the public.

I am troubled by this conclusion for two reasons. First, it places primary emphasis on standardized test performance rather than on instructional improvement. Second, it focuses attention on minimums and basic skills when we really should be concentrating on more challenging goals and higher-order skills in our schools and colleges if education is to be equal to the economic forces of concern to the public.

Teachers, teacher educators, state officials, and test developers have a common interest in seeing to it that the standardized measures used are sound and fair. That is the purpose of this conference. But I believe we also have a common challenge: to help schools and colleges earn greater public confidence in order to shift attention in the future from minimum skills toward instructional improvement. In this way, success in achieving higher educational standards may become a reality for all students and all teachers. That, too, should be a purpose of this conference.

With that overall perspective, I will address the four questions presented to the panelists. Because the organization I head develops and administers one particular teacher testing program, I will call upon this experience in answering the four questions. *It should be recognized that other organizations develop teacher tests, and some states develop their own. The issues raised by these questions, therefore, would still be present even if the ETS's National Teacher Examination (NTE) programs did not exist.*

MAJOR QUESTIONS

What Are the Appropriate Uses of Testing Within the Teaching Profession?

Teacher tests have a place in American education. Properly developed and validated, they can measure the academic knowledge of prospective

teachers at the point of entry into the profession. Forty-one states now have such a requirement for teacher certification. Within the limits of any standardized paper-and-pencil examination, teacher tests provide evidence that a prospective teacher (*a*) has a basic knowledge of the subject (or level) he or she plans to teach, (*b*) has the minimum pedagogical knowledge that experienced practitioners and teacher educators deem necessary for beginning teachers, and (*c*) demonstrates certain basic communication skills necessary to instruct children in an elementary or secondary school classroom. I believe it is appropriate to use successful performance on a teacher test as a condition for entry into the profession and so, too, by vote of its Representative Assembly in 1985, does the National Education Association.

Twenty-five states currently require testing for admission to teacher education programs. The concerns of colleges and of state legislatures about competition over scarce funds for higher education have prompted some states and some institutions to require students to demonstrate adequate knowledge of basic skills before they begin teacher preparation. Evidence of basic skills mastery before matriculation as a teacher education major also is now required by new standards of the National Council for Accreditation of Teacher Education. Such requirements at least enable students to discover in advance any weaknesses in basic skills that might present difficulties in qualifying for certification two years later—*before* they elect to enter a teacher preparation program and while they still have time to strengthen these skills.

I advise states that have such an admission requirement for teacher education to consider offering an alternative to students who, though not meeting the qualifying test standard, achieve a test score within a previously designated range of the qualifying standard. The alternative should enable an aspiring teacher to demonstrate promise to a faculty committee, based on other measures of academic performance in the freshman and sophomore years of college. Such an alternative requires, however, that the college or university provide supplementary help so that students admitted in this manner have a reasonable opportunity to succeed on subsequent certification tests.

Beyond these two early points in the professional continuum, I do not believe there should be mandatory testing of teachers. Once employed and in the classroom, a teacher should be judged primarily on the basis of systematic supervision and evaluation of actual teaching performance with children. I support the Carnegie Task Force on Teaching as a Profession in its recommendation for new, higher-level professional assessments for purposes of career advancement and recognition. The ETS

is committed to working with the Carnegie group and others to develop new approaches for measuring teacher knowledge and skills. Such assessments, however, are a professional choice for teachers to make individually and should not be mandated.

Who Should Control the Content, Construction, and Use of Tests?

I firmly believe that experienced classroom teachers should have a central role in determining policy and content for all tests that affect their chosen profession. Teachers have been represented on the NTE Policy Council since it was first established by the ETS in 1979. They will comprise fully half the membership of the new Teacher Programs Council being established this year to implement the ETS's broadened commitment to services for teachers, as approved in June 1986 by the ETS Board of Trustees. Teacher educators also should have a direct role in determining policy and content for teacher testing programs. The new Teacher Programs Council, like its predecessor NTE Policy Council, will include strong representation of teacher educators (along with school administrators, a chief state school officer, and public representatives).

Teachers and teacher educators should also have a central role at two other critical stages in test development. First, they should be on the test committees that determine test specifications and develop the questions for each edition of a teacher test. It is standing policy of the ETS that every NTE test committee must have such representation. Teacher members of the Policy Council and of all test committees are selected, as a matter of policy, in careful consultation with the National Education Association and the American Federation of Teachers to assure that both of the nation's leading teacher organizations are directly represented on these decision-making groups.

A second key point for involvement is in the state process for deciding whether or not to select a particular test for the certification of teachers who want to be employed in that state, and for determining the qualifying score for certification. The *Guidelines for Proper Use of NTE Tests*, published by the ETS, require that a validity study be conducted by any state that wishes to use NTE tests, and that teachers and teacher educators participate both in that validity study and in determining qualifying scores for certification.

The purpose of the validity study is to assure that test content appropriately reflects the teacher preparation curriculum in *that* state and that the test content is relevant to the professional job requirements of begin-

ning teachers in *that* state. Thus, teachers and teacher educators have a central role in determining the content of the national NTE program and the relevancy of this content in their specific states.

As long as there are tests, there will be debates about what content should be tested. *These decisions should not be made by the test maker but by consensus of experienced professionals who are teachers and teacher educators in the specialty field concerned.* This consensus should be reviewed, on a regular basis, as new editions of the tests are developed. This is the practice for all ETS-developed teacher tests. For instance, at the urging of NEA President Mary Futrell and others, the Professional Knowledge Test of the NTE Core Battery was reviewed and, as a result, is being revised this year by a committee of teachers and teacher educators.

Despite claims to the contrary, the experience of the ETS has been that there is more consensus on the role of teachers than most people believe. In 1984, we conducted an analysis of the professional functions of teachers. A sample survey involving eighteen hundred teachers in three states revealed a high degree of agreement about these professional functions (Rosenfeld, Skurnik, & Thornton, 1986).

Genuine involvement of teachers and teacher educators in determining policy and content for teacher tests requires close working relationships between test makers and professional associations. I am very grateful for the cooperation we have received from the National Education Association, the American Federation of Teachers, and the American Association of Colleges for Teacher Education. I believe that this collegiality and collaboration sets an appropriate example for professional assessment in the United States. The ETS and I are committed to continuing and enhancing this example in the future as we broaden ETS services to preservice and in-service teachers.

How Can Public Demands for Teacher Accountability Be Translated into Appropriate and Equitable Testing Policies and Procedures?

Two facts are clear regarding this question. The first is that most states already have responded to public demands for teacher accountability. That die is cast in law or regulation. Second, if there is no significant change in attraction to and the strength of teacher preparation by the year 2000, the percentage of minorities in the teaching force of the United States may be cut almost in half from its current level of approximately 12 percent. This decline will be taking place at the same time as

the proportion of minority students enrolled in American schools is increasing dramatically. As documented in two ETS research reports, this decline is in part the result of a high percentage of minority candidates failing to qualify on state teacher certification tests (Goertz, Ekstrom, & Coley, 1984; Goertz & Pitcher, 1985).

It is a natural and predictable reaction to blame lower minority test scores on test bias. Something encouraging is happening, however, that counsels against such a reaction. The performance of Black and Hispanic students on the SAT, College Board Achievement Tests, Advanced Placement tests, and on exercises in the National Assessment of Educational Progress (NAEP) is improving steadily. Similar improvement has been reported on state basic skills tests. And qualifying rates for Black candidates on the NTE Core Battery are improving faster than for White candidates. These minority students are demonstrating that they can and will do better on standardized tests if they have better educational opportunities.

While efforts must be continued to prevent bias in standardized tests for teachers (and students), priority also must be given to strengthening educational opportunities for those who aspire to teach. In one such example, the ETS has been working with the nation's Historically Black Colleges (HBC) to assist these institutions in their efforts to improve undergraduate education generally and teacher education in particular.[1] Some of the most promising reforms in teacher education that I have seen nationally are taking place in historically Black institutions. The ETS also is working with several Hispanic organizations on collaborative projects to help improve educational opportunities for Hispanic students.

Properly developed and validated, standardized tests can be useful barometers of the quality of education provided to minority and majority students. Such usefulness requires that the tests be as free of bias as is possible. I believe that the ETS is in the forefront of test development organizations in guarding against bias in the tests it develops. The committees of teachers and teacher educators that develop NTE test questions are biracial, as is the NTE Policy Council. Before the new NTE Core Battery was inaugurated in 1982, biracial panels of experienced classroom teachers chosen independently by the National Education Association and by the American Federation of Teachers examined results for every question in the field test of the new Core Battery. In addition, all ETS test questions go through a mandatory sensitivity review process in which specially trained ETS test development specialists examine each question for potential bias on the basis of race, sex, or ethnicity. The NTE and all other tests we develop must conform fully to the ETS

Standards for Quality and Fairness,[2] and compliance with these standards is audited annually by internal and external groups. The validity studies required of states that wish to use NTE tests provide an opportunity for them to review the tests for bias independently. In addition to this scrutiny by knowledgeable reviewers, the ETS introduced this year a new statistical procedure for examining group differences as part of the process of selecting questions for its teacher tests.[3]

Important as these safeguards are, at the heart of increasing minority representation in the teaching profession are (*a*) early identification of those who might want to be teachers, (*b*) opportunities for strong high school and undergraduate educations, (*c*) financial aid incentives for entering teacher education, and (*d*) making the teaching profession generally more attractive in the competition for all talented youth. Such improved opportunities are essential for the equity and excellence so necessary if the teaching profession is to thrive as it should.

As someone whose entire professional career has been committed to the cause of equal educational opportunity, I believe that we do not serve children well or fairly by giving them teachers who have not themselves mastered the basic skills that the children must learn before they can graduate from high school. If those who aspire to teach cannot qualify on state-required teacher tests, the solution is not to do away with the tests. The solution is to improve the education being provided to aspiring teachers. That is what is needed—not permitting inadequately prepared teachers, White or minority, to inadequately prepare children who need and deserve better in the classroom.

What Steps Can Be Taken to Prevent the Misuse of Teacher Tests?

My answers to the preceding questions include important principles that I believe should underlie the proper ***development*** of teacher tests. The more complex task is how to prevent misuse in ***decision making*** and ***interpretation*** by those who use the test results.

The basic need is to inform all concerned about what teacher tests ***cannot*** do. No standardized tests that I know of can accurately measure human qualities such as dedication, caring, perseverance, sensitivity, and integrity. Yet when we remember outstanding teachers from our childhood, these are the qualities that are memorable. Tests also present and measure only a sample of the knowledge required for teaching. While practitioners decide on this sample, experts may differ on the choices.

No test results can guarantee that a prospective teacher will become a really *good* teacher in the classroom. Certainly no teacher can be successful without a strong knowledge of the subject taught and of the skills of teaching, but professional performance requires more than academic knowledge. Passing the Medical Board exams does not assure that a physician will have a caring bedside manner, nor does passing the bar exam guarantee an attorney's integrity. However, such tests do provide the public with evidence that a candidate has at least the basic knowledge, deemed relevant by the profession, to perform important parts of the job. That is the limited purpose of professional certification and licensing.

The ETS makes a concerted effort to promote the proper use and interpretation of tests it develops. Each testing program has guidelines on proper test use that are provided to those who take the tests and those who require them. Unlike a government agency, the ETS has no authority to enforce its program guidelines. The Board of Trustees and I recognize, however, that the ETS does have influence. We are trying to become more effective in exercising that influence to promote proper test use. Over the last five years, the ETS has intensified its efforts—many of them centered on teacher testing—to promote proper test use through various means: persuasion, research, cooperation, guidelines, legal intervention, and withdrawal of services.

Persuasion

Proper test use has been a central theme in ETS-sponsored conferences, training workshops, publications, and media interviews. Most of my public presentations address proper test use and the importance of recognizing the limitations of tests. Since coming to the ETS in 1981, I have given more than 250 media interviews, most of which covered some aspect of proper test use. The ETS also has provided advice on test use to many public agencies and private organizations and at times has publicly opposed specific uses of tests when such opposition was warranted.

Research Reports and Access to Data

A major portion of the annual $15 million budget for continuing research at the ETS is targeted at improving the development, use, and interpretation of tests on a long-term basis. Specific priority—representing approximately $5 million in 1985-86—is given to research on the improvement of minority education. I believe that this commitment to research is basic to our role as a nonprofit, educational organization. In

addition to this research commitment, the ETS and program boards like the NTE Policy Council have broadened access to data from tests—for researchers and others—through public-use data tapes, technical reports, and other services.

Cooperation with Other Organizations

In addition to the HBC/ETS collaboration referred to above and continuing collaboration with the client boards that sponsor ETS-developed tests, the ETS has cooperated with a number of organizations to promote better understanding and use of tests. Among these groups are the American Association of Collegiate Registrars and Admissions Officers, the American Federation of Teachers, the Hispanic Higher Education Coalition, the National Education Association, the National PTA, the National Urban League, and the Southern Regional Education Board. We also are working with other test publishers and the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education to develop a "Code of Fair Testing in Education."

Program Guidelines

ETS-developed testing programs, including NTE programs, publish guidelines on proper test use and score interpretation. These guidelines are issued by the program sponsors of each test or by the ETS for test programs it sponsors directly. Extensive training, technical assistance, and other guidance are provided in support of these guidelines.

Legal Intervention

Periodically, a conflict over the use of tests is so serious that it raises legal questions that become the subject of a court suit. In a number of instances, the ETS has provided expert witnesses and has been prepared to submit *amicus curiae* briefs to provide the courts with its position on proper test use.

Withdrawal of Services

In three instances, the ETS has felt that the proposed use of a test was so improper that it must refuse to let its tests be used for such a purpose. In Texas and Arkansas, NTE testing services were not made available by the ETS to carry out new state laws that required on-the-job teachers to pass a test in order to remain in their chosen profession, regardless of

years of service and satisfactory evaluations by supervisors. We also withdrew services from the Houston Independent School District when it decided to use test results to freeze the salaries of experienced teachers. These instances represent the first time any test producer in the United States has taken such action. The ETS has shown that it will act, when circumstances warrant, to demonstrate its commitment to the proper use of tests it develops.

THE QUESTION THAT WASN'T ASKED

Where should we be heading in the future? I believe attention in educational reform should shift from an obsession with accountability toward a commitment to instructional improvement. Greater priority should be given to helping the in-service teacher succeed and grow.

- We need teacher centers that are readily available in school districts or collaboratives of school districts. These centers should be directed by teachers and oriented primarily to practical problem solving and instructional issues of concern to classroom teachers.
- We need equitable supervision and evaluation policies that are aimed at professional development and that involve teachers-helping-teachers.
- We need student assessment instruments that can help the student and teacher, on a continuing basis, with the diagnosis of where they are instructionally and how to progress to the next stage of learning.
- We need supports for the teacher like the new creative classroom testing software being developed jointly by the National Education Association and the ETS.
- We need self-assessment procedures so that interested teachers can seek to improve their skills under auspices of their professional association. The ETS is currently developing, for instance, an interactive videodisc assessment designed to help teachers improve their classroom management skills. Advances in technology now make possible an array of such supportive aids.
- We need greater involvement of teacher educators in the professional development of in-service teachers, and greater involvement of experienced teachers in preservice teacher education.

These are the kinds of initiatives that hold promise for improved achievement by students and by their teachers. The educational reform movement started because the public wanted children to learn more and better. This goal is more likely to be achieved by helping the teacher rather than by testing the teacher. It's time that we got on with that challenge.

Notes

[1]The HBC/ETS collaboration began in 1983. Under the direction of a steering committee of HBC presidents, the collaboration has sponsored joint workshops and other activities in such areas as financial aid, educational evaluation, teacher education, instructional use of computers, assessment of learning outcomes, and access to graduate and professional schools. Four workshops for faculty were conducted in summer 1986 in support of HBC efforts to improve teacher education.

[2]These standards are consistent with the 1985 *Standards for Educational and Psychological Tests* of the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education.

[3]This procedure is consistent with the resolution on testing adopted by the Board of Directors of the American Association of Colleges for Teacher Education at its July 1986 meeting.

References

Goertz, M.E.; Ekstrom, R.B.; and Coley, R.J. 1984. *The impact of state policy on entrance into the teaching profession.* NIE Grant No. 683-0073, Final Report submitted to the National Institute of Education. Princeton, N.J.: Educational Testing Service.

Goertz, M.E., and Pitcher, B. 1985. *The impact of NTE use by states on teacher selection.* Princeton, N.J.: Educational Testing Service.

Rosenfeld, M.; Skurnik, L.S.; and Thornton, R.F. 1986. *Analysis of the professional functions of teachers—Relationships between job functions and the NTE Core Battery.* Research Report 86-8. Princeton, N.J.: Educational Testing Service.

Sandefur, J.T. 1986. State assessment trends. *AACTE Briefs* 7 (August): 12.

* * *

Q. You speak of two objectives: to restore public confidence in education and to improve to the extent that we can the quality of the teaching force. Is it conceivable that these are inseparable entities? Or would we be smarter to attack these two objectives differently?

A. I think you cannot do one without the other. Public confidence is not going to be restored unless teachers are helped to improve learning. The public is not satisfied with how kids are doing, and it has some reasons to feel that way. Take a look at the new NAEP report on writing skills, on the inadequacy of that performance. The public knows that. Employers see it. The colleges see it. (I don't want to let colleges off the hook on that, either.) The key is not to ask who can we blame but to ask how can we help people so that the situation gets corrected, so that the public's confidence is increased? I think teachers and schools will benefit if we answer that question.

Q. You say that ETS is committed to working with the Carnegie group and others to develop new approaches for measuring teacher knowledge and skills; that such assessments, however, are a professional choice for teachers to make individually and should not be mandated. If we follow that dictate, how then are teachers accountable?

A. The nice thing about being a professional is that you feel a sense of accountability. I have not met many teachers who do not want to do a better job. They really want to do what's right for kids. They need help, and they want support; but they're willing to do it. The advances in technology are making possible today new kinds of assessment techniques that ten years ago were not possible. We all beat up on multiple-choice exams, but they're the best that we have. Now better is possible, and it is important that we get on to better. The classroom management skills videodisc that I referred to says, "We've had a panel of experienced teachers say these are the best ways to handle that situation. Why don't you think about this alternative, this approach, this concept?" That's the way you develop somebody professionally. You do not say, "Here's the right answer." There are now ways to assess situationally appropriate or correct responses—a very exciting and promising development for the teaching profession.

Q. You spoke eloquently about the great importance of seeing to it that teachers are well educated and well informed. Would it surprise you to learn that a teacher who got a rather low score on the NTE nonetheless got outstanding results in teaching learning disabled

students? If you were a principal of a school and had such a teacher, what would you do?

A. I would congratulate her, thank her, bless her. Again, the score is useful for purposes of certification—not hiring, but certification. Hiring goes on in the interview process. I used to be a principal. I did not look at test scores to decide whether I was going to hire a teacher. I did look to see if the teacher was certified. I was more interested in talking to that teacher, trying to get a sense of what kind of human being he or she was, what kind of work experience the teacher had, talking with a person who had worked with the teacher before, trying even to talk with a parent or maybe a former student to find out what that teacher was like in terms of inspiration. The test score is not and should not be the hiring decision. We have removed NTE from the local selection process for that very reason.

Q. You argue forcefully about the consummate importance of providing minority teachers with repeated opportunities to overcome inequities and inadequacies in their training. But we need to ensure that we have a representative minority membership in the teaching force. Given a cadre of minority teachers who fail to qualify on the tests, would we be better off to coach them to pass your tests or to go back and give them a good general education?

A. I hope you'll ask some of the other leaders from the Historically Black Colleges about how they are approaching that on their campuses. They have made a conscious decision, and I believe an absolutely correct one, not to try to teach to the test. It won't work. You need somebody in the classroom who has the skills that the test, on a sampling basis, attempts to measure. But you are not going to help teachers succeed if you just get them over the hurdle of the test. You have got to be sure they are educated more broadly than that.

Q. You are in the business of testing knowledge and you are now trying to improve teachers' classroom management skills. Isn't that a little greedy?

A. As a matter of fact, the work that we're doing related to instructional improvement is not likely to be able to support itself. So we have to earn enough in some other programs to be able to support those efforts. We will also be recommending that a major amount of our revenues be devoted to various public service activities. It is not greed. There is a need out there, and there is something that an organization like ETS can do to help teachers succeed.

TESTING TEACHERS: STRATEGIES FOR DAMAGE CONTROL

John G. Weiss, Executive Director
Center for Fair and Open Testing (FairTest)

Standardized tests have become America's cradle-to-grave arbiter of social mobility. The scores from over 40 million standardized multiple-choice tests annually help determine who will and who will not—

- Obtain a high school diploma in over two dozen states
- Gain admission to three thousand colleges, universities, and professional schools
- Be permitted to sell insurance or real estate, walk a police beat, fight a fire, practice law or nursing, run a golf course, frame a picture, or enter over one hundred other occupations.[1]

Nowhere is the rush toward testing more evident than in the teaching profession. Two decades ago, standardized tests for entering teachers were a rarity. Today, forty-five states mandate that entering and/or practicing teachers must pass standardized multiple-choice tests. A typical teacher now has to take up to a dozen different standardized tests—including high school graduation exams, the PSAT, ACT, SAT, achievement tests, PPST, the GRE, NTE, competency tests, and licensing exams. Yet there is little evidence that any of these exams accurately predict an individual's capacity to be an effective teacher. A recent National Institute of Education-funded study by Pan American University Professor Peter Garcia (1985) concluded that teacher tests "have no predictive validity for future performance."

Note: Unlike the other conference papers, this paper was written after the conference and submitted for inclusion in the proceedings. The paper reflects Mr. Weiss's presentation at the conference but also includes new material to which conference participants did not have an opportunity to respond.

There are other serious problems with these tests and the way they are used. George Madaus (1985), in his National Council on Measurement in Education presidential address last year, stated, "Our tests are being burdened with too much responsibility. Present and proposed uses [of tests] will diminish teacher judgments to a minor role and change the very structure of American education." In Madaus's view, "Today, tests must be just good enough to pass legal muster and cheap enough to be acceptable to budget conscious policy makers." Strong words for the president of an organization comprised largely of individuals involved in making tests, but nevertheless quite true.

THE IMPACT OF TESTING ON THE TEACHING PROFESSION

How will these barely adequate but just affordable standardized tests "change the very structure of American education"? In the next five years, a million new teachers must be hired. This means that 40 percent of those who will make up America's teaching force in 1990 are not yet teaching (Pressman & Gartner, 1986). But as a direct result of our heavy reliance on teacher-licensing tests, few of these new teachers will come from minority backgrounds.

For instance, only 15 percent of Black teacher candidates passed Alabama's certification test, as compared with 78 percent of the White candidates. In Florida, 92 percent of the Whites but just 37 percent of the Blacks passed that state's teacher certification exam. As a result, just 3.6 percent of Florida's new teachers will be Black. Standardized tests will also reduce Black representation in Arizona, California, and Texas to just 2 percent, in Connecticut to less than 1 percent, and in Oklahoma to 1.5 percent (Smith, 1986).

In 1980, minorities comprised 25 percent of our public school population and 12.5 percent of our teaching force. If current trends continue, by 1990 minority teachers will be just 5 percent of the teaching force (Pressman & Gartner, 1986) while 40 percent of America's public school population will come from minority groups (Smith, 1986).

These data suggest a deepening crisis. Not only do standardized tests systematically decrease access to an important profession for minority adults, but tests also deprive *all* students of minority role models during their formative years. As Walter Mercer has pointed out, "To operate a public school system without black teachers is to teach white supremacy without saying a word" (Pressman & Gartner, 1986, p. 12).

PEDAGOGICAL IMPLICATIONS OF TESTING TEACHERS

Our recent headlong rush to test teachers is also having a disturbing impact on what is being taught in our nation's teacher training programs. Schools of education, especially those dedicated to preparing minority students, are redesigning their curricula to conform to the limited content of what a standardized multiple-choice test can measure.

In 1979, the year in which Louisiana mandated that all teachers pass the Educational Testing Service's National Teacher Exam, only four of sixty-nine Grambling State University students passed the test. By 1983, fourteen of sixteen Grambling students—or 88 percent—passed the NTE. Grambling's efforts to boost its success rate included—

- The adoption of screening tests to deny low-scoring students admission to teacher training programs
- In-school coaching courses designed to prepare students to take the NTE
- A directive to all faculty members to teach to the test (Pressman & Gartner, 1986).

The University of Arkansas at Pine Bluff has also implemented "an ambitious, comprehensive effort to improve student [test] performance, which has produced an increase in the pass rate on the NTE from 42% in 1983 to 73% in 1984" (Pressman & Gartner, 1986). Grambling's and Pine Bluff's promotional materials stress the institutions' high pass rates. Pressure to attract students is causing many—if not most—other teacher training programs to follow suit.

Efforts to boost scores do not take place only inside schools of education. In the last decade a $100 million ancillary industry has sprung up devoted exclusively to coaching students to take tests. For example, the Stanley Kaplan Test Preparation Company, a subsidiary of the *Washington Post*, is vigorously promoting its expensive teacher test preparation courses. Test coaching puts low-income students in double jeopardy: not only are they unable to afford the advantages promised by coaching companies, but successful coaching increases the disparity between advantaged and disadvantaged groups even further.

Moreover, the pedagogical ramifications of both in-school and out-of-school coaching courses are distressing. Students are spending hundreds of dollars and hundreds of hours cramming for multiple-choice exams that will not necessarily make them better teachers. In fact, if anything, such courses may make students cynical about the teaching profession.

USING STUDENTS' TEST SCORES TO EVALUATE TEACHERS

If what is happening in St. Louis today becomes a national trend, teachers may soon be faced with a new type of test: either boost students' scores or risk getting the boot. During the past year, the St. Louis Board of Education became the first in the country to use student test scores as part of a district's teacher evaluation process. *Education Week*, September 17, 1986, reported that fifty teachers were given unsatisfactory ratings solely because their students did not meet the required scores on the California Achievement Test. The resulting furor from the St. Louis Teachers Union forced the Board of Education to rerate the fifty teachers as "satisfactory" and "modify" evaluation procedures so that the exam did not carry so much weight. However, the board still plans to use the California Achievement Test as one of four evaluation criteria, a decision that is currently being legally challenged by the Teachers Union, which claims the test was not designed for such a purpose. Perhaps the board should reflect on George Madaus's (1985) warning:

> When teachers' professional worth is estimated in terms of exam success, teachers will corrupt the skills and reduce them to the level of strategies in which the examinee is drilled.

Many educators are well aware of the danger to the curriculum represented by the misuse of standardized tests. For instance, Stanford University Professor Larry Cuban has warned:

> When the model curriculum standards are wired to the tests, what you get is measurement-driven instruction, a bureaucratic, systems management approach to teaching. When all these things come together, as they are coming together in California, Texas, and a few other states, the intent is to determine as precisely as possible what teachers will teach. . .that turns teachers into technocrats. . .the creative and imaginative part of teaching shrinks. (Dronka, 1986)

And as RAND Corporation's Linda Darling-Hammond (1984) and many others have documented, this regimentation of the school curriculum is causing many qualified and superb teachers to leave the profession.

THE UNACCOUNTABLE TESTING INDUSTRY

Public policy protects the food we eat, the products we buy, the safety standards at our workplaces. Yet, despite its overwhelming impact on both citizens and institutions, the testing industry is essentially unregu-

lated. Few consumer protection laws apply, and test publishers refuse to voluntarily provide elected government officials, independent researchers, and test takers themselves with information needed to verify that exams are fair and valid.

Both individuals who believe they have been wronged by a standardized test and the growing evaluation reform movement in this country have been seriously hampered by the inaccessibility of information. Without access to data, they cannot investigate abuses or design remedial procedures. The late Dr. Oscar Buros, editor of the authoritative *Mental Measurements Yearbook*, pointed out in 1977:

> Today it is practically impossible for a competent test technician or test consumer to make a thorough appraisal of the construction, validation, and use of standardized tests. . .because of the limited amount of trustworthy information supplied by test publishers. (Strenio, 1981, p. 274)

Unfortunately, during the past decade test publishers have become more secretive, deceptive, and sophisticated in the techniques they use to prevent their products from being scrutinized. Even federal government sources of information on tests are controlled by test publishers. For example, America's largest testing firm, the Educational Testing Service, operates the federal government's Educational Resources Information Center on Tests, Measurements, and Evaluation (ERIC/TME)—the nation's "official" clearinghouse distributing information on standardized tests and other methods of evaluating students and workers. FairTest has documented that this clearinghouse blatantly promotes ETS products and philosophy.[2]

WHAT CAN BE DONE?

Mounting evidence demonstrates that standardized multiple-choice tests often provide little indication of a person's ability and unfairly discriminate against Blacks, Hispanics, rural Americans, females, and unusually creative people of all backgrounds (Rosser, 1987; Garcia, 1985; Darling-Hammond, 1984). Despite these serious flaws, it is not likely that the number of standardized tests administered each year will decline in the near future. Until new forms of evaluation are developed, simplistic numerical scores from these exams will continue to influence the educational and employment opportunities of millions of Americans.

In the short run, five simple, inexpensive safeguards can help ameliorate the unfortunate consequences caused by teacher tests.

1. *There must be truth in testing.*

Test takers and researchers need access to information about teacher tests. Even raw statistics about different ethnic groups' pass rates are difficult to obtain. As University of Florida Professor G. Pritchy Smith (1986) explained:

> A Catch-22 scenario is often presented to the researcher. Testing companies most often insist that state departments of education, not they, have the responsibility for reporting test results. State departments, in turn, insist that since the testing company administers the test, it should report the results. The ultimate machination occurs, however, when both testing companies and state departments declare that it is not they but the colleges and universities that have the responsibility to report test results. (p. 3)

It is equally difficult for independent researchers to gain access to exam questions. Former National Council on Measurement in Education President George Madaus (1985) reports that "the simple truth is that many tests by vendors for state agencies are secret, and not readily available for peer evaluation." Indeed, except for university admission tests, which are covered by truth-in-testing laws, tests are not available for inspection by the individuals whose lives could be altered dramatically by the test results.

Under truth-in-testing provisions, students taking college and graduate school exams may obtain copies of the test questions, their answer sheets, and the "correct" answer key. In addition, complete copies of past exams are made available periodically to both students and independent researchers. This allows students to prepare for the exams. In addition, it gives researchers the ability to examine test companies' claims that their tests are valid and fair. Over one million students have had their scores raised due to errors discovered through New York's truth-in-testing law.

Only one teacher test is covered under New York's truth-in-testing law—the ETS Graduate Record Exam (GRE) in Education. This exam is used by many graduate schools of education across the country. The first and thus far only copy of the GRE-Education exam to be released illustrates the poor quality of an exam which passed all of the nation's most prominent test makers' "quality control" standards. It is a perfect example of that "barely adequate" quality criticized by George Madaus. Look at question 170, for example:

Which of the following would be the most appropriate intelligence test in assessing the intellectual functioning potential of a 6-year-old child who has limited exposure to educational resources and materials?

(*a*) Stanford-Binet Intelligence Scale

(*b*) Wechsler Intelligence Scale for Children

(*c*) Lorge-Thorndike Intelligence Test

(*d*) Otis-Lennon Mental Ability Test

(*e*) California Test of Mental Maturity[3]

The test makers apparently do not expect the students to be aware that Federal Judge Robert Peckham ruled in 1979 that the Wechsler Intelligence Scale for Children was unfairly biased against minority children, since (*b*) is given as the "correct" answer.

Only public exposure will force test manufacturers to reexamine items and improve their quality. That is why many education organizations join FairTest in supporting truth in testing. The American Association of Colleges for Teacher Education's Board of Directors recently recommended that truth in testing be extended to teacher tests. The NEA's New York affiliate has endorsed legislation which would extend the truth-in-testing principle to teacher tests. In fact, of the organizations represented at this conference, only ETS opposes initiatives to extend the truth-in-testing concept to teacher tests.

Examine the GRE-Education questions that the ETS has been forced to disclose (see Appendix), then decide for yourself why the ETS is against efforts to make additional tests public.

2. *Tests must be made as fair and valid as possible—the Golden Rule Bias Reduction Principle.*

Safeguards need to be established to ensure that teacher tests measure relevant knowledge differences between test takers and not irrelevant, culturally specific factors. One way to make tests as fair as possible is the Golden Rule procedure. This is an objective technique, based on a November 1984 out-of-court agreement between the Educational Testing Service, the state of Illinois, and the Golden Rule Insurance Company. The agreement settled a lawsuit charging that the ETS Illinois Insurance Agent Licensing Exam unfairly discriminated against Blacks and was not job-related.

The Golden Rule reform makes exams *fairer, not easier*. The principle requires that the same content areas be covered as in previous tests, and that the exam is at the same level of overall difficulty. The only differ-

ence is that, within groups of items measuring the same content area, the test publisher must select those that display the least difference in correct answer rates between minority and majority test takers. As Emory University Professor Martin Shapiro told the *New York Times* on November 29, 1984, "Once you have this method, to not use it is to knowingly use a more discriminatory procedure."

To understand why the Golden Rule reform is needed, it is necessary to examine how most test makers currently construct exams, including teacher licensing tests. For each content area, test publishers develop a pool of potential questions. They then pretest these questions on a sample group of test takers. Next, test publishers discard the items that they believe are ambiguous, biased, or otherwise flawed. From the remaining pool of items they employ a statistical technique—based on *r*-biserial correlations—to select for the final test those pretested questions that ***maximize*** differences between test takers. Often questions which maximize differences between high- and low-scoring students are really measuring test takers' knowledge of irrelevant, culturally specific information.[4] By using such items, tests discriminate against otherwise qualified individuals.

Recently, application of the Golden Rule procedure reduced the difference between the scores of White and Black candidates on ETS's Illinois Insurance Agent Exam by 25 percent. Since many minorities score just below the passing cutoff, a 25 percent reduction in the disparate impact of the exam will lead to a substantial increase in the number of minority candidates who pass—without making the test one iota easier (Shapiro, 1986). Statisticians on the Illinois Insurance Agent Exam advisory panel predict that the disparate impact will be reduced even further as the Golden Rule procedure is used on future editions of the test (Shapiro, 1986).

A November 1986 out-of-court settlement extended a modified version of the Golden Rule concept to Alabama's Initial Teacher Competency Test, published by National Evaluation Systems of Amherst, Massachusetts.[5] Legislators in California, Massachusetts, New York, Rhode Island, Texas, and Wisconsin are currently debating whether to extend the Golden Rule principle to exams administered in their states.[6]

To our dismay, several testing companies and their consultants have begun to distribute misleading information about FairTest's position on Golden Rule. FairTest does not believe, as ETS President Greg Anrig asserts, "that group differences on test questions primarily are caused by 'bias'" (Anrig, 1987; also see Weiss, 1987). Rather, FairTest recognizes that group score differences reflect a host of causes, including genuine

knowledge differences, test-taking abilities, as well as the inclusion on tests of irrelevant and biased questions. The purpose of the Golden Rule reform is to help ensure that biased test questions are removed from exams.

Another misinterpretation of FairTest's position on Golden Rule, made recently by University of Illinois Professor Robert Linn and National Evaluation Systems Attorney Michael Rubell, is that our goal is simply to extend the out-of-court settlement reached in Illinois (Rubell, 1986; Linn & Drasgow, 1987). FairTest has always recognized that the Illinois settlement has many special provisions that either were incorporated into the settlement to appease one of the parties in the case or are only appropriate for an insurance licensing exam.

To set the record straight, FairTest is simply working to extend the Golden Rule principle, which is that: *Among questions of equal difficulty and validity in each content area, questions which display the least differences in passing rates between majority and minority test takers should be used first.*

3. *Test publishers should be held responsible for how their tests are actually used.*

Standardized tests are like drugs: properly used they can be beneficial, but if misused they can cause enormous damage. A test designed for one purpose or group can have disastrous consequences if it is administered to a different group or used for an unvalidated purpose.

For decades, test publishers have warned that the results from a single administration of a test should never be used to make an important decision about individuals. Yet today, many teacher tests are used in just such a fashion. Test company officials attempt to deflect criticism by saying that other factors are used in conjunction with test scores. But their argument is often flawed. George Madaus (1985) has no doubt of the overriding importance of the test scores:

> There is one inescapable fact. When the test score is not a sufficient condition, but a necessary one for certification or promotion—when it can override all other indicators that may point to a different decision—then the score *is* an infallible arbiter. A single test score can stop you dead in your tracks.

Currently, ETS's Pre-Professional Skills Test (PPST) is stopping many minority students from even entering teacher training programs. Last year in Texas, only 23 percent of the Blacks and 34 percent of Hispanics passed this test, compared with 73 percent of the Whites (Pressman & Gartner, 1986). In Buffalo, New York, 25 percent of minority teacher

candidates and 76 percent of Whites passed the PPST the last time it was given, according to *Education Week*, February 11, 1986.

As the situation now stands, test publishers make the tests and then claim no responsibility for their misuses. When faced with a blatant abuse, they simply point to some manual stating that the test should not be used for that purpose. They think that is the end of their obligation. Increasingly, judges and legislators have stepped into the breach to compel responsible test use. The Golden Rule and Alabama settlements are just two examples. If manufacturers of standardized exams do not take affirmative action to stop misuses and abuses, we may soon see product liability regulations extended to the entire industry.

4. *There must be an independent ERIC.*

Researchers, test takers, policymakers, and the public need an independent source of information about the accuracy and cultural biases of these exams. While the federal government theoretically has just such an entity—the ERIC Clearinghouse on Tests, Measurements, and Evaluation—it is operated by America's largest testing company, the Educational Testing Service. Having ETS operate the ERIC/TME Clearinghouse is like having a fox guard the chicken coop. Secretary of Education William Bennett and Attorney General Edwin Meese should act on FairTest's petition charging that it is a conflict of interest for a company that earns nearly $200 million annually from selling tests to control the government's official clearinghouse about their products.[7]

5. *Test users and people concerned about the misuse of standardized tests should support FairTest and other organizations working to ensure that tests are fair, open, valid, and properly used.*

There is a tremendous amount of money to be made in standardized testing—well over a billion dollars a year. There is almost no money to be made if you are a test critic. Accordingly, the current debate has been one-sided. FairTest is working with researchers, educators, and civil rights activists around the country to balance the debate, but we can only do it with your help and your information.

America needs a vigorous debate about our uncritical rush toward testing teachers. This event, which brings together test manufacturers, test users, test takers, and test critics, should be considered a model for the kind of dialogue which we, as a matter of urgency, need in this country. The National Education Association deserves all our thanks for opening the dialogue.

APPENDIX

The GRE-Education Test

ETS's GRE-Education Test helps determine who will and who will not be admitted to many of America's most selective graduate schools of education. To date, only one form of the GRE-Education Test has been made public under New York State's truth-in-testing law.

Over thirty questions on this two-hundred-item exam seek information about standardized testing. No question hints that tests can be biased, inaccurate, or inappropriately used.

The test is overwhelmingly oriented toward the accomplishments of (mostly dead) White men. Questions seek information about the accomplishments of over sixty-seven White males, seven White females, one Black man, and no Black women. No questions seek information about the educational needs or accomplishments of Hispanic or Native American populations.

While ETS claims that this exam is not a "speed" test, students are given just fifty-one seconds to answer each question.

Examples:[8]

1. Most authorities consider intelligence to be influenced by which of the following

   (*a*) environment alone
   (*b*) primarily environment
   (*c*) heredity alone
   (*d*) primarily heredity
   (*e*) environment and heredity

36. Essay tests are not considered to be objective tests because

   (*a*) they have poor reliability
   (*b*) they cannot be machine scored
   (*c*) students must write out their answers
   (*d*) they often require expressions of opinion in the answers
   (*e*) scorers are likely to grade the answers differently

116. A high school counselor tells Harriett that she probably will do well in college because, "Based on the results of this scholastic aptitude test, you did better than 89 out of every 100 entering college freshmen." In order for the counselor to tell Harriett this, her test performance was probably reported as

   (*a*) a stanine
   (*b*) a percentile rank
   (*c*) a percentage score
   (*d*) an IQ score
   (*e*) A T-score

125. Criticisms of the public schools are most likely to arise during periods of

(*a*) rapid industrial progress
(*b*) concern and unrest about the future
(*c*) economic prosperity
(*d*) cultural lag
(*e*) crises faced by other countries

158. A well-known example of longitudinal research was concerned with gifted students identified by IQ tests. The person who initiated this research was

(*a*) Binet
(*b*) Guilford
(*c*) Terman
(*d*) L.L. Thurstone
(*e*) Wechsler

170. Which of the following would be the most appropriate intelligence test to use in assessing the intellectual functioning potential of a 6-year-old child who has had limited exposure to educational resources and materials?

(*a*) Stanford-Binet Intelligence Scale
(*b*) Wechsler Intelligence Scale for Children
(*c*) Lorge-Thorndike Intelligence Test
(*d*) Otis-Lennon Mental Ability Test
(*e*) California Test of Mental Maturity

The correct answers are as follows: 1. (*e*), 36. (*e*), 116. (*b*), 125. (*b*), 158. (*c*), 170. (*b*).

Notes

[1]A FairTest research project in progress has already documented that at least 40 million standardized multiple-choice tests are administered each year in America.

[2]This has been reported in the *Chronicle of Higher Education*, September 10, 1986. For more information, contact FairTest, Box 1272, Harvard Square Station, Cambridge, MA 02238.

[3]Reprinted by permission of the Educational Testing Services, the copyright owner.

[4]For example, examine the following item from the ETS's SAT:

RUNNER: MARATHON

(*a*) envoy:embassy
(*b*) martyr:massacre
(*c*) oarsman:regatta
(*d*) referee:tournament
(*e*) horse:stable

(The SAT question is reprinted by permission of the Educational Testing Service, the copyright owner.) Fifty-three percent of the Whites but just 22 percent of the Blacks gave the wanted answer (*c*). Twenty percent of the students would have obtained the correct answer if they had just guessed. Clearly this item does not measure students' "aptitude" or logical reasoning ability but knowledge of an upper-middle-class recreational activity. The Golden Rule reform would replace this item with another analogy question that did not require esoteric, culturally specific knowledge.

[5]For details of the *Allen v. Alabama State Board of Education* ruling, which challenged the constitutionality of a state teacher-testing program that had a disproportionate minority failure rate, see *Education Week*, December 10, 1986.

[6]For information on current Golden Rule initiatives, see *FairTest News Update*, No. 4.

[7]To obtain a copy of the FairTest petition, endorsed by researchers from over twenty states, send a stamped, self-addressed envelope to FairTest, Box 1272, Harvard Square Station, Cambridge, MA 02238.

[8]GRE questions are reprinted by permission of the Educational Testing Service, the copyright owner.

References

Anrig, G. 1987. Golden rule: Second thoughts. *APA Monitor*, Jan., 4.

Darling-Hammond, L. 1984 (July). *Beyond the Commission reports: The coming crisis in teaching*. Technical Report R-3177-RC. Santa Monica, Calif.: RAND Corporation.

Dronka, P. 1986. Writing model curricula. *Association for Supervision and Curriculum Development Update*, September.

Garcia, P. 1985 (August). *A study of teacher competency testing, and test validity with implications for minorities and the results and implications of the use of the Pre-Professional Skills Test (PPST) as a screening device for entrance into teacher education programs in Texas*. Report No. NIE-G-85-0004. Edinburg, Tex.: Pan American University.

Linn, R., and Drasgow, F. 1987. Implications of the golden rule agreement. *The Score* (American Psychological Association), Jan., 4.

Madaus, G. 1985 (April 2). Public policy and the testing profession—You've never had it so good. Presidential address delivered to the National Council on Measurement in Education annual meeting.

Pressman, H., and Gartner, A. 1986. The new racism in education. *Social Policy*, Summer.

Rosser, P. 1987. *Sex bias in college admissions tests: Why women lose out.* Cambridge, Mass.: National Center for Fair and Open Testing.

Rubell, M. 1986. Disparate impact of teacher competency testing on minorities: Don't blame the test makers or the tests. *Yale Law & Policy Review* 4: 375.

Shapiro, M. 1986 (December 11). Testing reform initiatives: 1987 and beyond. Paper presented at the annual convention of FairTest, Washington, D.C.

* * *

(Note: the following question-and-answer exchange was in response to Mr. Weiss's oral presentation at the conference. As previously noted, his paper was written after the conference and submitted for inclusion in the proceedings. While it reflects the presentation, it also includes new material to which participants did not have an opportunity to react.)

Q. By way of clarification, FairTest believes that good tests are a good thing, is that right?

A. If a test meets three criteria it should be given: it is properly constructed, it is properly and appropriately used, and it is properly interpreted. Oftentimes we're using the test without understanding the ramifications of what's going on. Then, too, we keep generating more and more tests.

Q. Do you think that in the instance where a test is flawed and has some bad items we could identify the bad items and fix them? Are they repairable?

A. Perhaps. One of the problems is that only 15 percent of the tests in the United States are covered by truth-in-testing. It is very difficult to get a copy of test questions. Most questions are secure, so until we have truth-in-testing regulations or resolutions or laws, we cannot investigate the testing companies' claims that their tests are fair.

Q. You seem to imply that, in many cases, poor tests are quite deliberate, that the test makers know better. Do you see this kind of thing as a plot?

A. I do not believe it is an intentional plot. We have only had standardized tests for sixty years. They were brought to this country largely by well-meaning people. It's just that we have not examined what's going on with standardized testing. It's very important to raise these questions. We do not have the answers, but we need to create a healthy, vigorous debate.

Q. You suggested that many of the test-development firms produce a poorly crafted test, limiting their expenditure in order to underbid a competitor. Would you say, then, that spending more money on preparation would result in better tests?

A. No. Throwing money at a problem won't solve anything. What we need to do is to create specific standards, say that independent researchers should examine test questions, and enforce these requirements.

Q. I understand that Secretary Bennett, acting on the advice of a panel, has now suggested that the federal government ought to be required to pay for state-by-state comparisons of results on tests. Would you favor or disfavor such comparisons?

A. There is a proper role for comparisons among states. Test scores have actually been very helpful to people wishing to show that a particular school district is not receiving its fair share of funds or that a school district has problems and needs extra money. The problem is how to keep states from competing against each other to boost their test scores. An example of this is the "wall chart" that arrays SAT and ACT scores across states. What it does is rank all the states according to the test scores. States with low rankings often feel pressure to employ artificial techniques to boost test scores. For example, Indiana and South Carolina officials have instituted programs to discourage low-scoring students from taking the SAT. While this will boost these states' average test scores, it will do precious little to improve the quality of education. Comparisons can put the wrong kind of pressure on the school systems, because the SAT was not designed for this purpose, as the College Board and ETS acknowledge.

THE APPROPRIATE ROLE OF TESTING IN THE TEACHING PROFESSION

Norene F. Daly, President
American Association of Colleges
for Teacher Education

The American Association of Colleges for Teacher Education is committed to the preparation of teachers who have the knowledge and skills necessary to meet the diverse learning needs of the children and youth whom they will one day teach. This commitment extends to holding the conviction that teacher education institutions must assume responsibility for assessing the individual's competence to teach before admission to the profession. The Association's position has been developed in a series of resolutions dating back to 1980 that have been adopted by its membership.

The AACTE believes that the issue of teacher competency testing cannot be treated in isolation, that is, outside of the broader context of competency definition and competency assessment in teacher education. The current practice of testing content knowledge and basic skills in isolation is, in most cases, removed from the frame of reference provided by the preparing institution's competency criteria.

COMPETENCY ASSESSMENT IN TEACHER EDUCATION

In 1980, AACTE institutional representatives adopted a resolution supporting an assessment of basic skills by the school, college, or department of education as a criterion for entry or continuance in teacher education. The Association's membership also adopted a companion resolution supporting assessment by the preparing institution as an exit requirement in teacher education (AACTE, 1986a). Clearly, these actions represent support for teacher competency testing; however, it should be noted that the resolutions call for such testing to be administered within the context of the total preparation program so as to ensure that competencies developed by the program are those measured by the assessment process.

49

In 1985, the Association reaffirmed this position in the report of the National Commission for Excellence in Teacher Education, which the Association cosponsored. The first recommendation contained in that report reads:

> In the campaign to attract more students into the teaching profession, quality must not be sacrificed. All teacher candidates should be held to rigorous standards and undergo at least three thorough reviews during the selection and training process. First, before admission into the teacher education program, every candidate should demonstrate above-average collegiate-level scholarship, good critical-thinking skills, and competence in communication skills, particularly reading, writing, and speaking. A variety of methods—paper-and-pencil tests, interviews, work samples, and the like—should be used to make these determinations. Second, before approval for student teaching, every candidate should demonstrate mastery of both the subject to be taught and the pedagogical foundations that underlie effective teaching. Third, before graduation from a teacher education program, every candidate should demonstrate his or her knowledge and skills on three measures: (*a*) a test of knowledge of the subject to be taught, (*b*) a test of knowledge and application of the foundations, science, and processes of teaching, and (*c*) ability to teach effectively.

In February 1981, the AACTE Board of Directors established two task forces that were charged with the task of defining competency attainment in teacher education. The Task Force on Profiles of Excellence was to develop a document that would describe what beginning teachers were to know and be able to do upon completion of their teacher education programs. The AACTE member institutions affirmed the position statement developed by this task force at the 1982 annual meeting. This affirmation led to the development of the Association's statement of essential competencies contained in *Educating a Profession: Profile of a Beginning Teacher* (AACTE, 1983b). The task of competency definition has continued since that time and is presently being pursued by two action groups operating within the Association's Center for Change in Teacher Education: the Knowlege Base Action Group and the Teacher Competence Action Group.

The second task force established in 1981 was asked to prepare a position statement on the competency assessment of beginning teachers. The work of the Task Force on Competency Assessment was then affirmed by the membership at the 1982 annual meeting when a four-part resolution was adopted that reaffirmed the membership's support for, and encouragement of, the development of competency assessment programs in member institutions; stressed the design of measures of attainment of appropriate professional skills, knowledge, and attitudes; called

for the development of programs in all teacher preparation institutions predicated upon clearly delineated criteria related to those skills, knowledge, and attitudes; stressed the fact that no program should be devised that places sole reliance upon a single measure or upon a single assessment technique; and called upon member institutions to seek the involvement of representatives from the professional education community in the preparation of assessment programs (AACTE, 1983a).

THE AACTE COMPETENCY ASSESSMENT MODEL

The position statement developed by the second task force, *Educating a Profession: Competency Assessment* (AACTE, 1983a), describes a model teacher education assessment program that could be put in place in member institutions to measure attainment of essential competencies by beginning teachers. The time sequence suggested for assessment of specified competencies would be as follows:

- *Prior to admission to teacher education.* At this point, basic oral and written communication skills and computational skills would be assessed using standardized or locally constructed evaluative instruments and/or writing samples, grades earned in liberal arts course work, interviews, and faculty recommendations. In addition, overall academic achievement would be assessed for formative, not summative, purposes. Grade point averages and/or ACT or SAT scores would also be used to develop this assessment.

- *Enrollment for each subsequent semester.* As the student progressed through the teacher education program, the achievement of essential competencies would be monitored by the student and the advisor. Measures that could be used in this process include anecdotal information provided by instructors and clinical supervisors, rating scales, and grades. The task force also recommended that these data be compiled in a student portfolio.

- *Prior to admission to student teaching.* The task force recommended that responsibility for assessing competency development prior to admission to student teaching be assumed by a committee rather than the student's advisor. Factors that would be considered at this point include cumulative data in the student's portfolio as well as letters of recommendation, anecdotal information, grade point average, and grades in specific professional courses.

- *Graduation from the program.* At this point, assessment would cover general knowledge, teaching field knowledge, professional knowledge, and teaching skills. The cumulative data compiled in the student's portfolio would be used, as well as audio and video tapes and student teaching records.
- *During the first year of employment.* The task force stressed the continuing role that teacher educators must play in the support and assessment of first-year teachers. This involvement has since been incorporated into the redesign of the standards of the National Council for Accreditation of Teacher Education (1986).

During the first years of employment, the AACTE believes that teacher educators, teachers, and administrators should be jointly engaged in the development and administration of an assessment model that will serve to meet the needs of the local school district as well as emphasize the continued development of those competencies that have been identified in the undergraduate teacher preparation program. The Association also believes that the competencies of beginning teachers should not be assessed using the same instruments and standards used with experienced professionals, and that the assessment process for beginning teachers should be administered in ways that will serve to identify means whereby the individual's strengths can be enhanced and deficiencies remediated.

Given the foregoing, there should be no doubt that AACTE believes in assessment as an entirely appropriate means whereby the essential competencies of beginning teachers can be measured. What should also be clear is the realization that the AACTE does not believe that any one measure, be it a paper-and-pencil test or data collected from clinical observations or other information, should be used to make admission or retention decisions at any point in the student's program or upon entry to the profession.

Since 1980, the AACTE has monitored the development of the teacher competency testing movement largely through the work of J. T. Sandefur (1984, 1985, 1986), an Association past president. In addition, Association institutional representatives in a number of states where teacher competency testing is mandated have been actively engaged in the development of such tests.

UNINTENDED CONSEQUENCES

There have been a number of unintended but nonetheless damaging consequences of the teacher competency testing movement.

First, at the national level the impact of competency testing on the availability of minority candidates for the profession has been severe to the point where it has now reached critical proportions.

Second, in those states where the candidate failure rate, both majority and minority, has been widely publicized, it has the potential to contribute to a serious lack of public confidence in the competence of teachers generally. In a sense, the movement could be viewed as providing the public the opportunity to confirm beliefs that both those teachers seeking admission to the profession and those already employed are not adequately qualified. Of course, the opposite may be true: The spread of teacher competency testing may enhance public confidence in the quality of the teaching force.

Third, a number of states have mandated cutoff scores for competency tests without benefit of adequate research information to support such decisions, thereby seriously undermining the confidence of education professionals in the ability of such tests to predict candidates' probable success or failure in classrooms. Indeed, a recent report of "rigorous new standards" to be enacted in one state called for admission to teacher education programs on the basis of a score at the thirty-fifth percentile or above on the communication and general knowledge tests of the National Teacher Examination (Olson, 1986).

Fourth, teacher education programs, in some instances, have been placed on probation or had program approval withdrawn on the basis of test results that may represent the performance of a small number of students, students who completed course work anywhere from five to fifteen years prior to taking the competency test, and/or students from out of state who completed only a small portion of their program at the institution cited as inadequate. Flippo (1986), Jacobson (1985), and others point to the dangers inherent in such practices, which ignore the fact that from 65 to 75 percent of the prospective teacher's preparation program takes place in arts and sciences departments and can have the effect of creating adversarial relationships among academic departments on the same campus, with the state department of education, and among universities competing for the best passage rate on competency tests. The use of teacher competency testing as an indicator of effectiveness of the institution's program rather than as an indicator of the candidate's competence represents a departure from the original purpose of such tests and could have the effect of undermining confidence in institutions identified as ineffective, thereby leading to a further diminution of the resource base for teacher education. This could have a particularly damaging impact on historically Black institutions.

Fifth, competition among institutions to successfully prepare students to pass teacher competency tests could lead to a "dumbing down" of the teacher education curriculum, where emphasis upon development of minimum competencies replaces attention to more significant and complex components of the curriculum. In some instances, teacher education institutions, in the interest of survival, have required that students pass the state competency test—intended as an assessment after program completion and prior to certification—prior to admission to student teaching. This practice would seem to indicate that the student teaching experience is not essential to the attainment of minimum competencies.

Sixth, the teacher shortage evident in certain regions of the country has been exacerbated by the widespread use of teacher competency tests as more and more candidates grow wary of seeking admission to a profession where public confidence and support have been eroded by the public "failure" of so many candidates. Indeed, teacher competency testing, originally intended to increase the status of teachers, may have had the effect of deterring many qualified candidates from taking such tests after completing their programs, thereby further diminishing the number of candidates in a particular field and opening the door to the issuance of emergency certificates or an increase in the number of those teaching out of field (Flippo, 1986). It may also happen that more able candidates will be attracted to the profession in greater numbers as they perceive the profession to have higher standards that are assessed through teacher competency testing.

Finally, in states where the competency tests in use are of insufficient rigor to adequately screen out under- or unqualified candidates, those persons are gaining access to the profession in greater numbers.

TEACHER COMPETENCY TESTING AND MINORITY CANDIDATES

While all of the foregoing consequences have the potential to further erode the teaching profession, few have more far-reaching implications than the minority teacher candidate issue. Teacher competency testing is not the sole reason for the critical shortage of minority teachers. Qualified minority candidates are being attracted to other professions and careers that heretofore were not open to them. However, the minority teacher issue is one that must be addressed by all who are concerned about the teaching profession and the preparation of well-qualified minority members for that profession.

In general, fewer minority students are attending college. Cooper reported that "from 1976 to 1981, the 119 historically black institutions experienced a loss of 13,000 full-time undergraduate student enrollment" (Givens, 1986). This phenomenon is also the reality on other types of campuses. Cooper also reported that "nationally, the comparable black-white success rate on the various standardized tests is approximately 86 percent for whites and 26 percent for blacks and other minorities" (Givens, 1986). The need to address the education of minorities in the K-12 sector and through transitional college course work seems apparent.

In July 1986, the AACTE Board of Directors adopted a resolution that recognized the prevalence of teacher competency testing and the need to safeguard against conditions that might have the effect of promoting discrimination based on sex, minority membership, or handicapping conditions as well as inhibiting the inclusion of minorities in teacher education and the teaching profession in proportion to their representation in society. It set forth the following conditions for the development and use of standardized tests used for admission to, or graduation from, teacher education programs and for the certification of teachers:

> Test developers should provide researchers and test users with statistical information about their test's reliability and validity. This information should include an analysis of different ethnic and gender groups' performance on the test as a whole, as well as on each item on the test.
>
> Truth-in-testing policies which include the disclosure of test items should be extended to these tests. Such policies should be modeled after those used for other standardized examinations such as the Scholastic Aptitude Test and Law School Admissions Test.
>
> Tests should be constructed so that, insofar as possible, individual or group differences in test scores result from real differences in people's knowledge of the subject being tested and not from inappropriate characteristics of the questions themselves. Whenever possible, and without lowering the validity or reliability of the test for any group, and without altering the difficulty level of items, tests should be constructed from items which display the least differences in passing rates between minority and majority examiners of comparable backgrounds and educational level.
>
> The American Psychological Association's *Standards for Educational and Psychological Tests* should be adhered to in the development of these tests. (AACTE, 1986b)

The ACCTE's position, therefore, is that tests should be considered as one measure of teacher competence and only if they meet the foregoing criteria. Such tests should be used, in combination with other measures, at the points in the preparation program cited above to assist teacher

educators in making important decisions as they and others attempt to address the minority candidate issue.

WHAT TESTS CAN AND CANNOT DO

The Association supports the use of tests to determine students' acquisition of basic skills. However, such tests must be viewed only as indicators that basic skills are, or are not, intact. Only direct observation of students in clinical settings will reveal whether students can apply basic skills as well as professional skills and other knowledge. The AACTE continues to support research efforts to determine what relationship, if any, exists between test scores and classroom performance. Pugach and Raths (1983), Weise and Harris (1984), and others have pointed to the lack of evidence that teacher competency tests are predictive of classroom performance.

The public demand for teacher accountability should be viewed as presenting an opportunity for teachers, teacher educators, test developers, and others to produce and present evidence that the teaching act is far more complex than can be measured by a test of basic skills and that such tests should not be considered as the sole cause for eliminating or choosing candidates.

THE CONTENT, CONSTRUCTION, AND USE OF TEACHER COMPETENCY TESTS

The content, construction, and use of tests as part of the total assessment program in teacher education should be controlled by representatives from the professional education community, i.e., those who have the professional knowledge and expertise to determine how such tests should be constructed, normed, and interpreted. These representatives should first be involved in determining the competencies to be assessed by such tests and the degree of competency attainment necessary to indicate proficiency.

The tests that result from this process should be subjected to rigorous validation procedures, with every effort possible made to ensure that they represent a consensus adequately reflecting the professional education community's judgment as to what beginning teachers must know and be able to do. Such tests must also build upon and reinforce other evaluative procedures in the total competency assessment model. It is only when such steps have been taken and all of the stakeholders in the assessment process feel confident that the tests have credibility that they

will begin to serve the purpose for which they were originally intended, that is, to ensure that only the best qualified candidates are admitted to the teaching profession.

References

American Association of Colleges for Teacher Education. 1983a. *Educating a profession: Competency assessment.* Washington, D.C.: American Association of Colleges for Teacher Education.

_____. 1983b. *Educating a profession: Profile of a beginning teacher.* Washington, D.C.: American Association of Colleges for Teacher Education.

_____. 1986a. AACTE: Resolutions. In *AACTE directory 1986-87.* Washington, D.C.: American Association of Colleges for Teacher Education.

_____. 1986b. Report from the Task Force on Testing to the AACTE Board of Directors. Unpublished memorandum. Washington, D.C.: American Association of Colleges for Teacher Education.

Flippo, R.F. 1986. Teacher certification testing: Perspectives and issues. *Journal of Teacher Education* 37 (No. 2): 2-9.

Givens, S. 1986. Black teacher educators probe impact of reform movement on black colleges. *AACTE Briefs* 7 (No. 9): 6-7.

Jacobson, R.L. 1985. States' efforts to improve schoolteaching called inadequate. *Chronicle of Higher Education*, March 6, 20.

National Commission for Excellence in Teacher Education. 1985. *A call for change in teacher education.* Washington, D.C.: American Association of Colleges for Teacher Education.

National Council for Accreditation of Teacher Education. 1986. *NCATE redesign.* Washington, D.C.: National Council for Accreditation of Teacher Education.

Olson, L. 1986. Panels propose teacher-training reforms. *Education Week*, Dec. 3, 12.

Pugach, M., and Raths, J. 1983. Testing teachers: Analysis and recommendations. *Journal of Teacher Education* 34 (No. 1): 37-43.

Sandefur, J.T. 1984. State assessment trends. *AACTE Briefs* 5 (March): 17-19.

_____. 1985. State assessment trends. *AACTE Briefs* 6 (March): 21-23.

_____. 1986. State assessment trends. *AACTE Briefs* 7 (March): 12-14.

Weise, K., and Harris, C. 1984. To test, or not to test. . .begging the question. *Teacher Education and Practice* 1 (No. 1): 17-19.

* * *

Q. You argued that the current practice of testing content knowledge and basic skills in isolation from the frame of reference provided by the preparing institution's competency criteria is a bad thing. If there is a disparity between what the tests are seeking to measure and the institution's criteria, who do you suppose is fallible? Is it that the institution has it wrong or that the test makers march to the beat of their own drum?

A. I think there is enough guilt to go around. I do not feel that teacher educators and teachers and test makers have worked together closely enough to develop competency assessment models. AACTE believes that in many cases those competencies that are being assessed are not necessarily those competencies that are developed in teacher education programs, and there is something radically wrong with that.

Q. At another point you say that the teacher educators must play a role in the continued support and assessment of first-year teachers, suggesting that once the teacher has entered practice the professor comes back and helps to assess. Do you think the profession would stand for that?

A. I not only think the profession would stand for that, I think the profession would encourage that. We have a model in the state of Oklahoma, the first-year entry program where teacher educators and teachers and administrators are very actively involved in the nurture and support of the first-year teacher. That model is rapidly gaining credibility around the country and has been replicated to some extent in other states. It is the only type of model that I think can be viable. Who knows better than the teacher educators who have been involved in the development of that teacher what should be done in order to support that teacher during the first year?

Q. You suggest that the candidate failure rate, both majority and minority, has been widely publicized, contributing to a lack of public confidence in the competence of teachers generally. And that would seem to suggest that if teachers have higher scores on tests, public confidence would be restored. During recent time we have also had heavy media coverage of the fact that students cannot write. Given the situation in which the nation's teachers score high on tests but the students still cannot read or write, would there then be a continued problem with respect to public confidence?

A. Keep in mind that when I talked about the diminution of public confidence in teachers as a result of the public failure of teachers, I also said that the opposite could be true—that the confidence of the public could be enhanced as a result of what it feels are rigorous measures to ascertain the skills of teachers. We have a way of blaming teachers for everything, which is totally out of line. There are so many factors involved in whether or not children have gained skills in school, factors having to do with the environment from whence those children come and the skills that the children bring to school. For me to say that there is a simple answer to your question would be simply to duck the issue. There is no simple answer.

Q. In your commentary on the tests per se, you seem to indicate that the tests are not doing what they are supposed to do. Do you feel that the tests seeking to measure teacher competency are too hard or too easy?

A. I do not think it is really an issue of whether the tests are too difficult or too easy. I think it is more an issue of what the tests are testing. What is the relevance of that to what teachers know, are able to do, and will be expected to do in the classroom? Again, I would call for a closer collaboration among test makers, teacher educators, and teachers to determine what it is that we should be testing and the most appropriate ways of testing. In addition, we do not know if there is a correlation between test scores and eventual classroom performance. If we screen candidates out of the classroom, which in effect is what we are doing, we will never know whether their performance could enhance the learning of youngsters. Given the fact that the teacher competency testing movement is a relatively new movement, research needs to be done to determine whether there is a correlation.

II. Reconsidering the Appropriate Role of Testing in the Teaching Profession

HIGHLIGHTING THE CRITICAL ISSUES

Louis J. Rubin
Professor of Education, University of Illinois

Surprisingly, there is a considerable degree of consensus in the positions of NEA, ETS, FairTest, and AACTE. All four presentors acknowledge a viable and legitimate role for testing. The agreement seems to verify the ancient adage that it is not crisis, but the way people respond to crisis, that determines survival. What we need at the moment is intelligent action.

Each of the speakers, for example, is concerned about the inhibiting effects of testing on minority representation. The hard facts are that present testing procedures work a grave hardship on minority teachers, and a better way must be found.

All the presentors also make it abundantly plain that standardized paper-and-pencil tests cannot accurately predict the quality of teacher performance. The problem with such tests is that they ignore qualities like creativeness, imagination, caring, as well as the crucial ability of teachers to interact skillfully with students.

There is also general agreement that teachers must play a role, perhaps a dominant role, in determining test content. The degree of involvement, however, is a point of contention. Sharon Robinson argues that teachers should govern both the content and the administration of professional testing. Norene Daly believes, on the other hand, that teacher assessment should be controlled by representatives from the professional education community, teachers included. Greg Anrig contends that experienced teachers should have a central role in determining test content and policy for all tests affecting the teaching profession. Drs. Daly and

---

Note: Dr. Rubin served as conference facilitator. This statement was presented to highlight the issues noted by the panelists for the benefit of conference participants, prior to their involvement in small group discussions.

Anrig are also concerned about public confidence. But by inference, I think we can assume that Sharon Robinson and John Weiss share their apprehension.

The solution, however, is anything but clear. Should we make the tests simple or difficult? Easy tests offer little protection. Hard tests produce higher failure rates, and often engender public alarm. What, then, would constitute a sensible position?

John Weiss is particularly worried about truth in testing, adhering to Golden Rule principles, ensuring that tests are job validated, eliminating sex bias, and controlling unwarranted comparisons. Above all, Mr. Weiss wants us to clarify the true effects of coaching on test scores.

Dr. Robinson bases much of her reasoning on the assumption that the distinctive knowledge base for teaching can be codified and tested. She is, to be sure, not alone in this assertion. Indeed, the National Board for Professional Teaching Standards and the 1986 Carnegie Task Force report reflect a parallel conviction. She points out, in addition, that current teacher tests do not adequately reflect this knowledge base, and that far better assessment procedures are possible.

Dr. Anrig, I suspect, would not deny that we have yet to reach the ultimate in test design. He notes, furthermore, that knowledgeable practitioners may still be poor performers and, conversely, that those with limited codified knowledge may nonetheless be particularly effective teachers.

There are also disagreements. Whereas Robinson and Anrig believe that teachers and district personnel should judge teaching competence after teachers enter service, Daly, in contrast, believes that teacher educators can and should play an important part in assessing and supporting the continued development of first-year teachers.

In another point of disagreement, Anrig contends that each state should conduct its own validity assessment to determine whether a teacher test is appropriate in that particular state. Daly argues that, instead, test developers should themselves provide researchers and test users with statistical information on a test's reliability and validity. Do states have responsibility for validating tests, or should the producers bear this obligation? If the burden lies with developers, organizations like ETS may need three, ten, or theoretically, fifty variations of the same test to ensure validity.

There are several other issues which we may wish to ponder. One has to do with premature action. If states make technical errors in the early stages of teacher competency testing—and we already have examples of tests that leave something to be desired—will the flaws create lasting

damage, or can we count on progressive evolution and improvement? Put another way, are states best advised to plunge ahead and do the best they can, despite our limited experience, understanding, technology? Or should the states delay matters until they are able to test intelligently? It's like the old joke about fruit-flavored wine. If you drink it you still get drunk, but the consolation is that at least you won't get scurvy. The immediate implementation of widespread teacher testing would demonstrate a willingness to police our ranks, but until the tests are beyond reproach, the cure may be worse than the disease.

Another concern relates to the impending teacher shortage. Secretaries Bennett and Weinberger recently suggested that we ought to utilize retired military personnel as classroom instructors. Neither of the major teacher organizations expressed opposition, the presumption being that some segments of the military have been educated at public expense, are knowledgeable in particular subject areas, have spent their service careers working with people and could, without too much difficulty, be retrained and put into the teaching ranks. They would, however, lack the broad background in pedagogy that we are beginning to associate with the genuinely professional teacher. How, then, should such "retooled" personnel be tested? What, in short, is the trade-off between high standards and overcoming what could be a critical teacher shortage?

The portent of differentiated staffing gives rise to yet another issue. It is becoming more and more clear that different subjects and different students require highly specialized techniques. We can no longer get by with teaching math and English in the same way. It may thus become politically and economically advantageous to assign teachers rather specific instructional tasks instead of expecting them to teach all things to all students. If that should occur, would we need to alter our approach to testing? For example, if we want a teacher who is adept at working with slow-learning children, and who excels at instilling a desire to learn, should we test for these particular capabilities or continue to test for total mastery in all areas of teaching? Pediatricians and psychiatrists, as an illustration, do not take the same Medical Board examinations.

Similarly, some teachers do extremely well in particular aspects of instruction. There are those who are very good at presenting information and providing explanations, those who are especially gifted in leading discussions, and still others who have a flair for guiding skill development. Should we then evaluate teachers' performance capability with respect to specific aptitudes, or should we take the position that every teacher who serves in the nation's schools must demonstrate complete mastery of the multiple things teachers are called upon to do?

Teaching, theorists have long reminded us, involves many different kinds of skills. There is didactic teaching, where the teacher explicates and clarifies, making the abstruse comprehensible. There is heuristic teaching, through which learners gain insight, find meaning, and learn to make sense out of their world. And there is yet another form of pedagogy in which teachers who are wonderfully proficient at asking probing questions help students to develop deeper understanding, think more astutely, and see unsuspected truth. Given these distinctions, and our growing awareness of stylistic differences in learning, what kind of teacher tests would make the most sense?

One final issue may also warrant attention—the matter of teacher morale. Three out of five teachers who begin their teaching career this fall will leave teaching within five years. The best and the brightest, moreover, are likely to be the first out of the door. This is not to say that the strong leave and the weak stay. It is to say, rather, that a high percentage of able, gifted teachers find teaching an unsatisfying vocation and choose, instead, to become flight attendants or personnel directors. Many teachers, in fact, discourage their own children from entering the profession. It would be the height of folly, consequently, to ignore the dangers of alienation and the effects of competency testing on professional spirit. If testing restores public confidence in schools but causes teachers to feel that they have once again been put upon, singled out, and unfairly penalized, we may win the battle and lose the war. What, in sum, can be done about controlling the conditions and context of testing so that further erosion of professional well-being does not occur?

PARTICIPANTS' OBSERVATIONS AND RECOMMENDATIONS

In small group discussions, the conference participants generated observations, points of consensus, and recommendations related to the appropriate role of testing in the teaching profession. These have been synthesized from the separate group reports and are listed below by topic area.

HOW, WHEN, AND WHAT TO TEST

- As a first step, we must decide what it is that teachers should know and be able to do. That is, we must define the body of knowledge teachers ought to master, who ought to have that knowledge—both among teachers and among teacher educators—and how that knowledge ought to be measured. And we need constantly to revisit any definition we develop, because it will change over time.
- At present, there is no body of knowledge that everyone agrees beginning teachers should know. Yet to a large extent, a body of knowledge is already defined by law and by the teacher preparation curriculum. Ultimately, any test supported by the profession must reflect a knowledge base that the profession can accept.
- We need to make current tests as fair and as valid as the state of the art will permit. As important, we need to develop alternative assessment procedures that will enable minority teachers to demonstrate that they have the essential knowledge and skills required of an educator.
- We need to set out the main purposes of testing in the profession. Will teachers and their students be best served by tests that are screening devices or by diagnostic instruments that could improve the quality of instruction?

- Instruments designed to measure pedagogical knowledge should be constructed so that they differentiate between those who have and those who have not received professional training.
- The profession needs to explore new means of assessing candidates' ability to select from their repertoire those skills and strategies that are appropriate in a particular instructional setting—new means such as technologies, novel test items, teaching vignettes, simulations, and so on.
- Testing has an appropriate role prior to licensure but not after licensure, with one exception. That exception is when an individual wants to change or to add to his or her area of certification. If testing is appropriate for initial certification, it is also appropriate to test those seeking certification in additional subjects or fields.

TESTING'S CIRCUMSPECT ROLE

- We must draw clear distinctions among testing, assessment, and evaluation.
- Testing should be one of several measures in a larger program that assesses individual performance prior to licensure. It is this larger program of assessment that might reasonably be expected to predict job performance.
- While paper-and-pencil, multiple-choice instruments that measure knowledge must have content validity, we should not expect candidates' scores on such tests to predict their teaching performance.
- It is inappropriate to use a test as the *only* criterion either for admission to a preparation program, for admission to the profession, or for on-the-job decisions about teachers. A test should be used as one of several criteria, including grade point average, observations, and interviews. Also, it seems reasonable to balance the criteria so that, for example, a high grade point average could compensate for a lower test score.
- No critical decisions, particularly employment decisions, should be based solely on test information. Practicing teachers must play a more prominent role, and indeed a major role, in determining who enters the classroom and who stays there.

- Nontest indicators, such as grade point average, may ultimately serve as a substitute for test scores if we strengthen those indicators by increasing academic expectations and providing quality instruction.
- Tests can and should be used to support teachers' professional development. Sometimes, however, tests are used in place of meaningful professional development, rendering professional development inadequate or missing altogether.

TESTING, PROFESSIONAL PREPARATION, AND TRAINING

- Questions need to be addressed about the quality of higher education in general and teacher education in particular. Discussions of teacher education should focus not only on the criteria for entry into teacher preparation programs but on the teacher education curriculum and the skills required of teacher educators.
- Undergirding the teacher testing movement is a basic concern with the quality of college preparation. We need to make university presidents, chancellors, boards of regents, and others aware of the fact that students' lack of basic reading, writing, and mathematics skills is a campuswide rather than specifically a college of education problem.
- We need to reach down into the K-12 grades to improve students' performance, because it is there that students receive a grounding in the foundations of learning.
- There is a clear and growing need for preservice and in-service training of teachers and administrators in the purposes, uses, and limitations of tests.
- All prospective teachers should attend an accredited teacher education program to ensure that their preparation meets standards established by the profession through the National Council for Accreditation of Teacher Education.
- Mandated testing is reasonably well established. However, we are entering a new era in which change and improvement cannot be mandated but must be conducted by the profession itself. The profession must come to agreement about teaching's knowledge base and adjust the teacher education curriculum in terms of the pedagogical skills and academic proficiencies expected of teachers. We have to improve teacher education in a way that will forestall additional testing mandates.

TESTING AND MINORITIES' ACCESS TO TEACHING

- All tests should be constructed to minimize racial differentiation.
- The profession must take steps to ensure that minorities become teachers and remain teachers.
- Minorities' differential performance on teacher competency tests relates to access to knowledge and test-taking skills, as well as to the structure of the test items.

DEVELOPING AN ACTION AGENDA

- There should be no additional test requirements, additional standards, or alterations in existing standards until it can be proved that high test scores for teachers may be related to quality classroom teaching. We need to collect data on the relationship between effective practice and test scores. We may discover there is no relationship.
- Students and teachers will need to pass currently mandated competency tests while the profession works to alter public testing policy. But we should not let testing dominate the curriculum as we teach teachers and students to pass the tests. Test taking should be taught as a separate skill.
- We should set higher standards within the profession and then advertise those higher standards so that strong new candidates will be attracted to teaching.
- The profession should set standards for its own assessment, advocate and demand multiple and alternative measures to demonstrate skill and knowledge, and demonstrate its willingness to be held accountable for its performance on these measures.

MOVING AN ACTION AGENDA FORWARD

- *Recommendation*: That a national validation study be undertaken to define the knowledge, attitudes, and values required of a beginning educator. This study should involve the education profession, including teachers, teacher educators, administrators, and chief state school officers. Results from the study should inform the development of appropriate assessment programs within the profession.

- *Recommendation*: That the profession identify existing assessment programs that it feels meet high technical and professional standards, and advocate the use and replication of these programs.
- *Recommendation*: That the results of this conference be disseminated widely to demonstrate the profession's concern with testing and to inform the development of appropriate testing practices and procedures for teachers.
- *Recommendation*: That this conference be viewed as a first step in defining the appropriate role of testing within teaching. From here we should move aggressively forward in a broad-based coalition of education organizations that will work with legislators, state boards of education, and others to alter public testing policy.
- *Recommendation*: That as a follow-up to this conference, a broader base of education organizations convene to define the competencies teachers are expected to possess so that the profession can then identify the characteristics of an appropriate assessment process.
- *Recommendation*: That research funds be made available to examine testing's impact on the supply of minority teachers.

PANELISTS' REACTIONS TO PARTICIPANTS' OBSERVATIONS AND RECOMMENDATIONS

Following oral reports from the conference discussion groups, the panelists from AACTE, ETS, FairTest, and NEA offered the following reactions and summary comments.

REACTIONS FROM NORENE DALY, AACTE

Among the points I heard was a general consensus against the use of tests as a single measure, that tests should be part of an assessment program. I also want to wrestle a little bit with the point about the distinction among testing, evaluation, and assessment, and testing as an artificial measure. If we agree with that, then testing can only be one component of an assessment model and evaluation of an individual is made on the basis of a total assessment package. That's a distinction worth keeping in mind.

I heard us agree about the need to exercise some political clout. We have come together in a group, we have had an opportunity to interact with individuals whom we have not heretofore had an opportunity to meet and talk to, and we have formed new alliances. It is important for us to keep those alliances alive and to exercise the clout that we can exercise, not as organizations alone, but as a coalition of NEA, AACTE, ETS, and FairTest. That's political clout.

There seems to be universal agreement that professional educators should exercise control over the standards-setting process and entry to the profession. It is important, then, for us to wrest control from those who see themselves in the driver's seat today.

The minority issue came up over and over again, and a recommendation that we must address that issue. It will not become any better if we ignore it. I think we have to address it now.

I think we have all agreed that testing has its place, that accountability models have their place. Teachers have to be accountable and teacher

educators have to be accountable. Now we have to determine the best means for assessing and holding those individuals accountable.

I liked the point of using test results to stimulate growth, not as a means of deterring individuals from entering the profession.

I heard over and over again the need to identify the competencies of teachers, and I would like to emphasize the need to identify the competencies of ***beginning*** teachers. I think we make a very serious mistake in the profession to hold beginning teachers to the same standards as we hold teachers who have had a good deal of experience. We have got to realize that when a teacher enters the profession, he or she is ready to do minimal things and not necessarily ready to do them at the very competent level teachers may achieve in five or six years.

Another point that came up over and over again was the elimination of test bias and, of course, that's tied to the minority issue. I was happy to hear reference to NCATE, because I feel that NCATE represents a model we could emulate. It represents the coalition of organizations around the issue of accreditation of teacher education. I think we could develop the same model to coalesce around the issue of accountability or assessment in teacher education.

I heard only one or two references to the issue of teacher education being a campuswide activity. I hoped there would be more attention given to the issue of the use of tests to close down programs, because that is something that has very serious implications for teacher education generally. In the state of Florida, for example, if 80 percent of an institution's candidates do not pass the teacher certification test—cumulatively over a five-year period—the programs are put on probation. We have had instance after instance where programs that are not producing many teachers have been placed on probation on the basis of the performance of one person on the test. We are analyzing data right now at our institution from the last teacher certification test. We have individuals who have come from other states; we have individuals who have taken one course in professional education. There are all sorts of abuses of the practice of using the results of tests to close down teacher education programs.

REACTIONS FROM JOHN WEISS, FAIRTEST

The first thing I learned from today's discussion is that it is very important that educators act and not react and that educators take the initiative. One reason we do not currently have multiple criteria in

assessment for teacher certification is that the profession was not as organized a few years ago as it is now in calling for multiple criteria.

There is real need for concern. You can see the writing on the wall. What is happening in St. Louis will be happening soon in your states and districts—pressure for using student test scores to hold teachers accountable. I personally believe that would be terrible. If you agree with me, get resolutions passed—in NEA, in AACTE, in ETS, in the American Psychological Association. Try to create a groundswell so that people fighting this kind of test misuse have tangible support.

Second, we are dealing with a crisis related to minority access to the teaching profession. I would advocate some scenario where we do a fifty-fifty weighting on GPA and test scores. That is not exactly what I would like, but it is a way to downplay the importance of test scores, and it may be politically acceptable.

Third, it is important that we make tests as fair as possible and eliminate cultural bias. We also need to have independent researchers examine the examiners. We need to be able to question what they are doing. It is not that all test makers are doing something wrong all the time; but we need to be able to examine their activities, to hold them accountable.

Last, I think we need to keep ourselves accountable. I would be disappointed if we came to the same conference next year and talked about the same issues with nothing changed. We need to make sure that what has happened here gets translated into action, and action needs to be taken quickly, because the world is changing quickly. We cannot talk only among ourselves. I would like to see a conference next year with the same four groups and one additional group—state legislators.

REACTIONS FROM GREGORY ANRIG, ETS

First, the point about grade point average. I like the idea of a sliding scale. You need to know, though, that the credibility of the GPA is quite suspect in the eyes of legislators. While the best predictor of freshman performance is the GPA from four years of high school, the best predictor for graduate education is not the GPA, it is the Graduate Record Exam. So there is a vulnerability that you have got to know about if you are heading down that road.

Second, painful as it may be, the issue of minority access to teaching is not going to be solved by tinkering with the test. I am not against tinkering, and I want to do anything we can do to make tests better. But when all is said and done, test performance is not going to change unless

the educational opportunities available to minorities change. We cannot kid ourselves about that. The longer we do, the more pressure we're taking off the system to change educational opportunities. Yes, the test has got to be scrutinized; yes, it ought to be attacked and challenged; where it can be changed, I want to change it.

Finally, I sat here listening to people describing available teacher tests, and since the National Teacher Examination happens to be the most widely used test, I was a little sensitive to that discussion. One of the things I take considerable pride in is that we always involve teachers and teacher educators in determining the content and the specification of the NTE. What you are saying ought to happen has happened. On the one hand, the concerns I heard about practitioner input suggest that it is not enough just to be represented. There has to be a feeling of ownership. On the other hand, 1.8 million members of the NEA cannot sit at the table and develop a test. Thirty-three hundred members of the AACTE cannot sit at the table and develop a test. You have to delegate that responsibility and gain ownership as a result of that involvement. If you do not do that, then all the talk today is not going to get us very far.

REACTIONS FROM SHARON ROBINSON, NEA

First, I want to talk about a notion that came out of every group and that makes me nervous: the need to validate the knowledge base for teaching, to determine what should be tested. That was good news. But I heard further that we need to involve everybody in the task. I thought about what that meant and decided that we probably could get it done in Yankee Stadium in about two years worth of time if we involve everybody who has an interest or represents an interest group. Everyone cannot and should not be involved or we will never get the task done.

Second, I heard a lot about linking the test to practice. To that end, I would ask us to begin to think anew about what it is we are trying to predict. A test in teaching ought to promise one thing and one thing only: that an individual has mastered what we have agreed to be the profession's requisite knowledge and skills. That is all. The test should not promise that the person will practice well.

Third, I do not think that we really understand what building a profession means. If we did, we would be thinking in terms of building a culture for success. We would not be worried about how many folks pass and how many do not pass. We would understand that the measure of our success is the measure of those who pass. Similarly, we would not be so concerned about identifying items that discriminate along the bell-

shape curve, because we would be building a system, an enterprise in which we would have virtually 100 percent success. So much success, in fact, that there would be room and time to go back and assist those who did not succeed.

We have to start inculcating in those who teach in classrooms today and those who will soon join them the notion that professional practice means building a culture of success, certainly for minority students, indeed for all students. And we know how to build that culture. All we have to do is get on with getting it done. Maybe then, maybe this time next year, we will be talking about a different set of questions.

I certainly would not expect us to launch into an agenda to professionalize teaching and overlook lessons from others' experiences. There are strategies we can adopt that would guard against the elimination of minorities from the teaching work force if we are determined that that will not be an outcome of professionalization. In other professions such disenfranchisement was an unstated intent, or no one was moved to stop it even though it was apparent.

I also think that, if we do not professionalize teaching by vesting in teachers the responsibility to implement professional practice on behalf of their students, we do not stand a prayer of seeing better education afforded minority students. What those students need, what all students need, is an environment for learning that is responsible and responsive to them, not one which is controlled by the state adoption of textbooks, by the district curriculum plan which says that at eleven o'clock on Wednesdays you're supposed to be working on math whether the students are ready or not. These regulations and structures have been put in place so that you have safe practice rather than professional practice. And by safe practice I mean standard practice. That is what we now have in place of a degree of confidence that will allow, encourage, even hold teachers responsible for doing what is needed to meet the educational needs of students. To get support from parents and the public we are going to have to say that teachers know how to do something important on behalf of their students.

EPILOGUE

Louis J. Rubin

If I had not been privileged to attend this event and instead read a newspaper story about the proceedings, I would have found the story heartwarming. I think the story would have demonstrated that the responses of the profession to the testing issue were nondefensive. The responses indicated the education profession's willingness to police its own ranks. The discussions testified to the fact that testing is a political if not a moral imperative. The educators present demonstrated a conspicuous willingness to acknowledge their own weaknesses and a willingness to seek better things.

Above all, what went on seemed to create a strong reaffirmation of the education profession's pedagogical creed and its beliefs. The people present seemed to demonstrate that they will not yield to the irrational, the illegitimate, or the ill-conceived. All in all, that made for a very good day's work.

CONFERENCE PARTICIPANTS

Robert A. Altman
Vice President
Educational Testing Service
Princeton, New Jersey

C. Leonard Anderson
Vice President, Region 1
Oregon Education Association
Portland, Oregon

Gregory R. Anrig
President
Educational Testing Service
Princeton, New Jersey

Tomas A. Arciniega
President
California State College
Bakersfield, California

Carl A. Backman
Dean
College of Education
University of West Florida
Pensacola, Florida

Eva L. Baker
Director
Center on Student Testing,
Evaluation and Standards
UCLA Graduate School of Education
Los Angeles, California

Leonard C. Beckum
Dean of Education
The City College of the
City University of New York

Herman E. Behling, Jr.
Assistant State Superintendent
Certification and Accreditation
Maryland State Department of
Education
Baltimore, Maryland

Terry Bergeson
President
Washington Education Association
Federal Way, Washington

Bruce M. Berman
Attorney-at-Law
Wilmer, Cutler & Pickering
Washington, D.C.

Henry Bern
Professor of Education
Indiana University
Bloomington, Indiana

Ermalee Boice
Assistant Executive Director
for Programs
Texas State Teachers Association
Austin, Texas

76

Ines A. Bosworth
Field Marketing Representative
Educational Testing Service
Princeton, New Jersey

Gerry Bracey
Director
Research and Evaluation
Cherry Creek School Division
Englewood, Colorado

Roxanne Bradshaw
Secretary-Treasurer
National Education Association
Washington, D.C.

Pat A. Browne
Teacher
Lewis B. Russell, Jr. School 48
Indianapolis, Indiana

Robert Calfee
Professor
School of Education
Stanford University
Stanford, California

Don Cameron
Executive Director
National Education Association
Washington, D.C.

Gene Campbell
Dean
College of Education
University of Arkansas
Little Rock, Arkansas

Norma Cantu
Education Director
Mexican American Legal Defense and
Educational Fund
San Antonio, Texas

Denise S. Carty-Bennia
Professor of Law
Northeastern University School of Law
Boston, Massachusetts

Jules Chametsky
Professor of English
Director
Institute for Advanced Studies
in Humanities
University of Massachusetts
Amherst, Massachusetts

Vernon L. Clarke
Dean of Education
Coppin State College
Baltimore, Maryland

Beverly P. Cole
Director
Education Department
NAACP Special Contributions Fund
Baltimore, Maryland

Nancy Cole
Dean
College of Education
University of Illinois
Urbana, Illinois

Jeremiah Collins
Attorney-at-Law
Bredhoff and Kaiser
Washington, D.C.

Fred Comer
Executive Director
Iowa State Education Association
Des Moines, Iowa

James Connerton
Executive Director
New Jersey Education Association
Trenton, New Jersey

Constance C. Cooper
Dean
College of Education
University of Akron
Akron, Ohio

William Corbett
Principal
J. R. Lowell School
Watertown, Massachusetts

James Crouse
Professor
School of Education
University of Delaware
Newark, Delaware

Margaret G. Dabney
Dean
School of Education
Virginia State University
Petersburg, Virginia

Norene F. Daly
President
American Association of Colleges
for Teacher Education
Washington, D.C.

Linda Darling-Hammond
Director
Education and Human Resources Program
The RAND Corporation
Washington, D.C.

S. John Davis
Superintendent of Public Instruction
Virginia Department of Education
Richmond, Virginia

Terry G. Davis
Attorney-at-Law
Seay & Davis
Montgomery, Alabama

Harold Dent
Chair
Testing Committee
Association of Black Psychologists
Berkeley, California

Penelope M. Earley
Director of Policy Development,
Public and Governmental Relations
American Association of Colleges
for Teacher Education
Washington, D.C.

Ronald Early
Instruction and Professional Development
Consultant
Florida Teaching Profession–NEA
Tallahassee, Florida

Robert L. Egbert
George W. Holmes Professor
of Education
University of Nebraska
Lincoln, Nebraska

Eugene E. Eubanks
Dean
School of Education
University of Missouri
Kansas City, Missouri

Mario Fantini
Dean
School of Education
University of Massachusetts
Amherst, Massachusetts

Joan First
Executive Director
National Coalition of Advocates
for Children
Boston, Massachusetts

Patricia S. Fleming
Professional Staff Member
U.S. House of Representatives
Subcommittee on Intergovernmental Relations and Human Resources
Washington, D.C.

Anne Flowers
Assistant Vice Chancellor for Academic Affairs
Board of Regents of the University System of Georgia
Atlanta, Georgia

Ralph Flynn
Executive Director
California Teachers Association
Burlingame, California

Henry T. Frierson
Associate Professor
University of North Carolina
Chapel Hill, North Carolina

Peter A. Garcia
Dean
Extended Education and Governmental Relations
Pan American University
Edinburg, Texas

William E. Gardner
Dean
College of Education
University of Minnesota
Minneapolis, Minnesota

Karen Garr
President
North Carolina Association of Educators
Raleigh, North Carolina

Warlene D. Gary
Associate Director
Government Relations
National Education Association
Washington, D.C.

Vincent C. Gazzetta
Director
Teacher Certification
New York State Education Department
Albany, New York

Keith Geiger
Vice President
National Education Association
Washington, D.C.

Beverly Glenn
Program Development Specialist
National Education Association
Washington, D.C.

Barry Goldstein
Assistant Counsel
NAACP Legal Defense Fund
Washington, D.C.

Donna Gollnick
Deputy Executive Director
National Council for Accreditation of Teacher Education
Washington, D.C.

Joe Grant
Executive Director
South Carolina Education Association
Columbia, South Carolina

Bert F. Green
Professor
Department of Psychology
Johns Hopkins University
Baltimore, Maryland

Henry J. Grubb
Assistant Professor
Psychology Department
East Tennessee State University
Johnson City, Tennessee

Walt Haney
Director
Center for Testing and Evaluation
and Educational Policy
Boston College
Chestnut Hill, Massachusetts

Catherine Havrilesky
Director
Teacher Programs and Services
Educational Testing Service
Princeton, New Jersey

Frances Haywood
President
United Teachers–Los Angeles/NEA
Los Angeles, California

George E. Hein
Dean
Advanced Graduate Study and Research
Lesley College
Cambridge, Massachusetts

Joan Herman
Associate Director
Center for the Study of Evaluation
UCLA Graduate School of Education
Los Angeles, California

Asa G. Hilliard
Fuller E. Callaway Professor
Georgia State University
Atlanta, Georgia

Blair Horner
Program Analyst
New York Public Interest
Research Group
Albany, New York

Paul Houts
Editor in Chief
Congressional Budget Office
Washington, D.C.

Willie T. Howard
Dean
School of Education
Howard University
Washington, D.C.

Paul R. Hubbert
Executive Secretary
Alabama Education Association
Montgomery, Alabama

David G. Imig
Executive Director
American Association of Colleges
for Teacher Education
Washington, D.C.

Mari Luci Jamarillo
Vice President for Student Affairs
University of New Mexico
Albuquerque, New Mexico

Nancy Jewell
Chair, NEA Standing Committee
on Instruction and Professional
Development
Oklahoma Education Association
Oklahoma City, Oklahoma

Burnett Joiner
Executive Academic Dean
Dean, College of Education
and Graduate Studies
Grambling State University
Grambling, Louisiana

Al Kaufman
Attorney-at-Law
Mexican American Legal Defense and
Educational Fund
San Antonio, Texas

Patricia M. Kay
Professor
Bernard M. Baruch College
New York, New York

Stephen L. Koffler
Program Administrator
Educational Testing Service
Princeton, New Jersey

Francine Korman
Associate Director
Instruction and Training
New Jersey Education Association
Trenton, New Jersey

Richard C. Kunkel
Executive Director
National Council for Accreditation
of Teacher Education
Washington, D.C.

Richard Lansford
Teacher
Jefferson County Public Schools
Lakewood, Colorado

Jim Lewis
Executive Committee Member
National Education Association
Washington, D.C.

Robert L. Linn
Professor, Educational Psychology
and Psychology
Department of Educational Psychology
University of Illinois at Urbana-
Champaign

Mary Ann McClean
Legislative Director
Office of State Senator
Kenneth LaValle
Albany, New York

Phyllis McClure
NAACP Legal Defense Fund
Washington, D.C.

Robert McClure
Director
Mastery In Learning Project
National Education Association
Washington, D.C.

James D. McCullough
Superintendent
Chattanooga Public Schools
Chattanooga, Tennessee

Page McCullough
Director
Atlantic Center for Research
in Education
Durham, North Carolina

David McDonald
Assistant to the Commissioner
Kansas Department of Education
Topeka, Kansas

Pearl Mack
Executive Committee Member
National Education Association
Washington, D.C.

George Madaus
Director
Center for the Study of Testing,
Education and Educational Policy
School of Education
Boston College
Chestnut Hills, Massachusetts

W. Frank Masters
Director
Research Services
National Education Association
Washington, D.C.

Marjorie M. Mastie
Consultant
Measurement and Guidance
Washtenaw Intermediate School District
Ann Arbor, Michigan

Donald M. Medley
Professor
School of Education
University of Virginia
Charlottesville, Virginia

Howard D. Mehlinger
Dean
School of Education
Indiana University
Bloomington, Indiana

Ernest (Pete) Middleton
Associate Professor of Education
University of Kentucky
Lexington, Kentucky

Marilyn Monahan
President
NEA-New Hampshire
Concord, New Hampshire

Damon P. Moore
President
Indiana State Teachers Association
Indianapolis, Indiana

Frank B. Murray
Dean
College of Education
University of Delaware
Newark, Delaware

Allan Nairn
Testing Project Coordinator
Center for Responsive Law
New York, New York

Carol Norman
Research Specialist
National Education Association
Washington, D.C.

Roger Olstad
Professor of Science Education
University of Washington
Seattle, Washington

James Outtz
Consultant
Outtz and Associates
Washington, D.C.

David Owen
Contributing Editor
Atlantic Monthly
Boston, Massachusetts

Rod Paige
Dean
School of Education and Behavioral Sciences
Texas Southern University
Houston, Texas

Claire L. Pelton
Teacher
Los Altos High School
Los Altos Hills, California

Vito Perrone
Vice President
Carnegie Foundation for the Advancement of Teaching
Princeton, New Jersey

William F. Pierce
Executive Director
Council of Chief State School Officers
Washington, D.C.

Marjorie Pike
Supervisor of Instruction
Robertson County School District
Springfield, Tennessee

John Poggio
Professor
School of Education
University of Kansas
Lawrence, Kansas

Paul Pottinger
President
New Manhattan Corporation
New York, New York

Robert Pressman
Attorney-at-Law
Center for Law and Education
Cambridge, Massachusetts

Diana Pullin
Associate Dean
School of Education
Michigan State University
East Lansing, Michigan

Nancy L. Quisenberry
Associate Dean of Academic Affairs
College of Education
Southern Illinois University
Carbondale, Illinois

Joe L. Reed
Associate Executive Secretary
Alabama Education Association
Montgomery, Alabama

Diane L. Reinhard
Dean
College of Human Resources and Education
West Virginia University
Morgantown, West Virginia

Roger L. Rice
Executive Director
Multicultural Education Training and Advocacy (META, Inc.)
Harvard University
Cambridge, Massachusetts

Priscilla Robinson
Reading Specialist
Eleanor Roosevelt High School
Greenbelt, Maryland

Sharon P. Robinson
Director
Instruction and Professional Development
National Education Association
Washington, D.C.

William L. Robinson
Director
Lawyers Committee for Civil Rights Under Law
Washington, D.C.

Louis Rubin
Professor of Education
University of Illinois
Champaign, Illinois

Lawrence M. Rudner
Senior Associate
U.S. Department of Education
Washington, D.C.

Philip J. Rusche
Dean
College of Education
University of Toledo
Toledo, Ohio

Gene Russinoff
Staff Director
New York Public Interest Research Group
New York, New York

John Ryor
Executive Director
Florida Teaching Profession-NEA
Tallahassee, Florida

Terry S. Salinger
Assistant Professor
Department of Teacher Education
University of Texas at El Paso

J. T. Sandefur
Dean
College of Education
Western Kentucky University
Bowling Green, Kentucky

Robert L. Saunders
Dean
College of Education
Memphis State University
Memphis, Tennessee

Dale P. Scannell
Dean
College of Education
University of Maryland
College Park, Maryland

Robert A. Schaeffer
Communications Director
Center for Fair and Open Testing
Cambridge, Massachusetts

Ana-Maria Schuhmann
Assistant Dean
School of Education, Technology
and Related Professions
Kean College of New Jersey
Union, New Jersey

Henrietta Schwartz
Dean
School of Education
San Francisco State University
San Francisco, California

Ramsay Selden
Director
State Education Assessment Center
Council of Chief State School Officers
Washington, D.C.

Rick Seymour
Staff Attorney
Lawyers Committee for Civil Rights
Under Law
Washington, D.C.

Martin Shapiro
Professor of Psychology
Emory University
Atlanta, Georgia

Theodore M. Shaw
Attorney-at-Law
NAACP Legal Defense Fund
New York, New York

Lorrie A. Shepard
Associate Professor
School of Education
University of Colorado
Boulder, Colorado

Thomas A. Shipka
Professor and Chair
Department of Philosophy
and Religious Studies
Youngstown State University
Youngstown, Ohio

Lee Shulman
Professor
School of Education
Stanford University
Stanford, California

Mike Slutsky
Attorney-at-Law
Cotton, Watt, Jones, King & Bowlus
Chicago, Illinois

Carol E. Smith
Senior Associate
American Association of Colleges
for Teacher Education
Washington, D.C.

David C. Smith
Dean
College of Education
University of Florida
Gainesville, Florida

G. Pritchy Smith
Chair
Division of Curriculum and Instruction
College of Education and Human Resources
University of North Florida
Jacksonville, Florida

Carl Stedman
Dean
College of Education
Austin Peay State University
Clarksville, Tennessee

Chuck Stone
Professor
Department of English
University of Delaware
Newark, Delaware

James H. Sutton
Director
Instruction and Professional Development
Iowa State Education Association
Des Moines, Iowa

Carol K. Tittle
Executive Officer
Ph.D. Program in Educational Psychology
Graduate School
City University of New York

John Trasvina
Legislative Attorney
Mexican American Legal Defense
and Educational Fund
Washington, D.C.

Marc Tucker
Executive Director
The Carnegie Forum on Education
and the Economy
Washington, D.C.

Shirley M. Turnage
Director of Pupil Services
Elzabeth City-Pasquotank
Public Schools
Elizabeth City, North Carolina

Dennis Van Roekel
President
Arizona Education Association
Phoenix, Arizona

Stanford von Mayrhauser
General Counsel
Educational Testing Service
Princeton, New Jersey

Mary Wakeford
Program Assistant
General Administration
University of North Carolina
Chapel Hill, North Carolina

Donald Watkins
Attorney-at-Law
Montgomery, Alabama

Gary Watts
Assistant Executive Director
Professional and Organizational
Development
National Education Association
Washington, D.C.

Paul Weckstein
Director
Washington Office
Center for Law and Education
Washington, D.C.

John Weiss
Executive Director
Center for Fair and Open Testing
Cambridge, Massachusetts

W. Scott Westerman
Dean
School of Education
Eastern Michigan University
Ypsilanti, Michigan

David White
Attorney-at-Law
Berkeley, California

Howard P. Willens
Attorney-at-Law
Wilmer, Cutler & Pickering
Washington, D.C.

Charlie G. Williams
State Superintendent of Education
South Carolina Department of Education
Columbia, South Carolina

James L. Williamson
Dean
College of Education
East Texas State University
Commerce, Texas

Charles Willie
Professor
School of Education
Harvard University
Cambridge, Massachusetts

Don Wilson
President
Ohio Education Association
Columbus, Ohio

Reginald Wilson
Director
Office of Minority Concerns
American Council on Education
Washington, D.C.

Paul Winkler
Superintendent
Lower Camden County Regional
High School District 1
Atco, New Jersey

Arthur E. Wise
Director
Center for the Study of the
Teaching Profession
The RAND Corporation
Washington, D.C.

Elaine P. Witty
Dean
School of Education
Norfolk State University
Norfolk, Virginia

Gerald Work
Professor
College of Education
University of Maine
Orono, Maine

Sam J. Yarger
Dean
School of Education
University of Wisconsin
Milwaukee, Wisconsin

Michael Zieky
Senior Examiner
Test Development
Educational Testing Service
Princeton, New Jersey